Western States Petroleum Association
Credible Solutions • Responsive Service • Since 1907

September 24, 2004

Via email to ddrechsl@arb.ca.gov

Dr. Deborah Drechsler, Ph.D.
Research Division, P.O. Box 2815,
California Air Resources Board,
Sacramento, CA 95814

RE: Benefits Analysis for the Proposed California Ozone Standard

Dear Dr. Drechsler:

The Western States Petroleum Association (WSPA) and the American Petroleum Association (API) represent companies that explore, develop, refine, market and distribute petroleum and petroleum products throughout the United States. WSPA is composed of nearly 30 companies that have operations within the 6 Western States including California. API is the national trade association of America's oil and natural gas industry, and represents more than 400 members involved in all aspects of the oil and natural gas industry. Both WSPA and API members are owners and operators of major facilities regulated under the California ozone standard, and producers and marketers of fuels that are often targeted as a means to reduce ozone precursors. As such, we have a direct and substantial stake on the outcome of this proposal. With this letter, we are providing comments on the draft benefits analysis (Chapter 10 of the Review of the California Ambient Air Quality Standard for Ozone) that has been developed by the California Air Resources Board.

We have asked Dr. Allen Lefohn of ASL and Associates and Dr. Stanley Hayes of Environ International Corporation to help us review Chapter 10 and prepare detailed comments. These nationally recognized experts have been deeply involved in many aspects of ozone research, have authored important sections of current and past chapters for the federal Ozone Criteria Documents or conducted research cited in previous Ozone Staff Papers, and helped develop and apply approaches for assessing ozone risk.

1415 L Street, Suite 600, Sacramento, California 95814
(916) 444-9981 • FAX: (916) 444-5745 • www.wspa.org

Our review has found significant problems in the methodology and assumptions that are used in the benefits assessment. We feel that the document should not be finalized until these problems are resolved This chapter moves far beyond previous health risk assessments conducted by EPA in their 1997 Ozone Staff paper, and introduces the concept of a link between ozone and mortality, which is likely an artifact of the analysis methodology and statistics. As such, we recommend that any benefits estimates predicated on mortality be removed, as these are not scientifically supportable.

It is critical that Staff address the issues raised herein, and a second draft of the Staff Paper be circulated for public review and comment. The most significant problems with the benefits analysis fall into the two areas outlined below. These problems are also described in more detail in our attached comments.

1. Inappropriate Choice of Rollback Model and Background Ozone Level

One of the most important aspects of the O_3 health benefits analysis is the adequacy of the rollback methodology. If the methodology is not adequate, then the application of the rollback model will introduce a level of uncertainty that may make the health benefits analysis unreliable. We believe this has happened. Staff has underestimated *policy-relevant background* at many areas across California by subjectively selecting a value of 0.04 ppm. Our analysis shows that the hourly average concentrations of *policy-relevant background* near sea level and at higher elevations will be higher at many locations than estimated by Staff.

In addition Staff concluded that the rate of change in the concentrations *above background* was similar among the percentiles and this observation justified its application of a constant percentage rollback to all sites within an air basin. In carefully reviewing the information in the Appendix for the rate of change in the concentrations *above background* from the 1980s, the data do not show a constant percentage change *above background* for each of the percentiles for the sites provided in the Appendix. Also, the rollback model used in the health benefits analysis does not necessarily mimic the slowing down of the mid-level hourly average concentrations in relation to the higher hourly average values and therefore, the predicted distribution of the concentrations may not be reliable.

2. Accrual of Health Benefits Below the Proposed Standards

The proposed California standards are based primarily on chamber studies involving human subjects. The chamber studies do not show statistically significant effects below the 8-hour, 0.08 ppm level. This led Staff to propose a standard of 0.07 ppm, to include a margin of safety. Note that in our prior comments, we cited research

by Dr. William Adams from UC Davis that does not support Staff's recommendations for a 0.07 ppm standard. This was based on analyses reflecting the inappropriateness of Staff choices for background level and the use of averaging exposures for the form of multi-hour standard.

In Chapter 10 no rationale is provided for accumulating health benefits below the level of the standard to an assumed *policy-relevant background* level of 0.04 ppm. Since approximately 76% - 86% of the benefits presented in Chapter 10 are accrued at O_3 concentrations between the proposed standards (0.07 ppm 8-hour and 0.09 ppm 1-hour) and background, accumulating benefits below the level of the standard enormously exaggerates the benefit of the proposed standard**.**

In addition, for more than twenty years, response to ozone has been directly measured in controlled human chamber studies. Ozone exposures in those studies have been characterized, exercise levels have been specified, controls to avoid confounding factors have been applied, and response indicators have been measured. Collectively, this body of human chamber studies provides a large and robust database with which to characterize human response to ozone over a wide range of ozone levels and lung function, respiratory symptom, and other endpoints.

However, none of these data are used in the benefits assessment in Chapter 10. Instead, Chapter 10 relies entirely on the results of epidemiological studies, bypassing data from human chamber studies. We strongly recommend that Chapter 10 be revised to include those chamber data in its assessment and to reconcile those data with assumptions made in the chapter based on epidemiological studies

Staff appears to have accepted epidemiological studies, many done for other purposes (e.g., PM) as being sufficient to establish causality and quantify concentration-response relationships for ozone. In fact, the large uncertainties and inconsistencies in this literature, documented by Staff and us, preclude interpreting and using the current epidemiological data in this fashion – particularly for ozone. For reasons not explained, the inconsistencies and model dependency of epidemiologically derived effect estimates – although recognized by Staff – are often subsequently ignored. It appears that there are greatly different standards of evidence that are being used in standard setting versus benefits estimation.

WSPA and API Recommendations

We do not believe that the benefits estimates as currently presented in Chapter 10 are based on sound scientific data. The estimates do not (1) use the correct body of data (chamber studies), (2) reflect the correct interpretation of the data used (e.g., assumes that acute mortality is causally linked to ambient ozone exposure), (3) use an appropriate estimation methodology (accrue benefits below the level of the standard and rely on linear concentration-response models), or (4) reflect the range of uncertainty associated with the data (potential unresolved confounding). As such, we fail to see

how the benefit estimates will help inform policy makers with their decisions on the proposed California ozone standard.

The chapter should be revised to address the significant issues raised in our comments before being presented to either the Air Quality Advisory Committee or the Board. In addition, to the extent that Staff or the Board feels it is necessary to have a benefits assessment performed, additional estimates should also be presented to the Board. At a minimum, benefit estimates should be based on models derived from chamber studies. One set of estimates should address the incremental benefits of achieving the proposed California ozone standard as opposed to meeting the current federal ozone standard. Estimates are also needed for the benefits accrued by meeting the proposed standard without accumulating benefits to the level of policy relevant background. All estimates based on epidemiological studies should incorporate the full range of uncertainty stemming from these studies. We recommend, however, that Staff remove any benefits estimates predicated on mortality, as these are likely artifacts of the analysis methodology.

Lastly, since the basis for the proposed California standard is identical to that used by EPA in their 1997 ozone Staff Paper – that is, human exposure effects demonstrated from chamber studies -- we believe that Staff should estimate ozone's impact by performing a risk assessment using a methodology similar to that used in the 1997 EPA Staff Paper.

We welcome the opportunity to continue discussions with your agency. After you have had a chance to review these submittals, please feel free to contact me at 310-808-2149, Mr. Kyle Isakower (API) at 202-682-8314, or Dr. Mark Saperstein, (BP, Chair of WSPA Task Force) at 714-228-6716.

Sincerely,

Michael D. Wang
Manager, WSPA

Cc: Dr. Alan Lloyd
Ms. Catherine Witherspoon
Mr. Mike Schieble
Ms. Catherine Reheis-Boyd

1115 11th Street, Suite 150, Sacramento, California 95814
(916) 444-9981 • FAX: (916) 444-8997 • www.wspa.org

Comments on the California Ambient Air Quality Standard for Ozone Document (CAAQSOD) Quantifying the Health Benefits of Reducing Ozone Exposure (Chapter 10)

Allen S. Lefohn, Ph.D.
A.S.L. & Associates
111 North Last Chance Gulch
Suite 4A
Helena, Montana 59601
406-443-3389
asl@attglobal.net
asl-associates.com

Prepared for

American Petroleum Institute
1220 "L" St., N.W.
Washington, D.C. 20005

and

Western States Petroleum Association
1415 L. Street
Suite 600
Sacramento, CA 95814

September 23, 2004

S. Summary

S.1 Introduction

The staff of the Air Resources Board (ARB) and Office of Environmental Health Hazard Assessment (OEHHA) has attempted to quantify in Chapter 10 of the California Ambient Air Quality Standard for Ozone Document (CAAQSOD) the adverse health effects of current O_3 levels in California by estimating the health benefits that would accrue from a hypothetical control strategy that achieves the proposed ambient air quality standards for O_3. In estimating the health benefits associated with reductions in levels of ambient O_3, the Staff has included the following four elements:

1. Estimates of the changes in O_3 concentrations due to a hypothetical control strategy;
2. Estimates of the number of people exposed to O_3;
3. Baseline incidence of the adverse health outcomes associated with O_3;
4. Concentration-response (CR) functions that link changes in O_3 concentrations with changes in the incidence of adverse health effects. These functions produce a beta coefficient, indicating the percent reduction in a given health outcome due to a unit change in O_3.

For performing the health benefits analysis in Chapter 10, Staff has relied on epidemiological results, even though Staff notes that there are a variety of unresolved statistical issues associated with the use of epidemiological results. Several challenges and unresolved issues present themselves with respect to designing and interpreting time-series studies of O_3-related health effects. Using the epidemiology data, the staff has focused on premature mortality, hospital admissions for respiratory diseases, emergency-room visits for asthma, school absences, and minor restricted activity days. For the purposes of its analysis, Staff has estimated health benefits down to an assumed background concentration of 0.04

ppm (except for emergency room visits for asthma, for which a higher threshold value was used). Approximately 76% - 86% of the benefits presented in Chapter 10 are accrued at O_3 concentrations between the proposed standards (0.07 ppm 8-hour and 0.09 ppm 1-hour) and background.

In this evaluation, a thorough review has been undertaken to evaluate the Staff's decision to (1) use 0.04 ppm as the concentration for *policy-relevant background*, (2) apply a proportional linear rollback model, and (3) substitute unreliable epidemiological results for the chamber analyses used to propose the level of the two standards.

S.2 The Selection of a *Policy-Relevant Background* of 0.04 ppm

One of the most important aspects of the O_3 health benefits analysis is the adequacy of the rollback methodology. If the methodology is not adequate, then the application of the rollback model will introduce a level of uncertainty that may make the health benefits analysis unreliable. For estimating the daily reductions in current O_3 concentrations that result at all monitoring sites due to a hypothetical control strategy, rollback factors from the 1-hour and 8-hour O_3 design values to the applicable standard were calculated for each air basin. The amount of rollback and the rate of change of the daily maximum 8-hour concentrations are both sensitive to the selection of the *policy-relevant background level*. A background O_3 concentration of 0.04 ppm (i.e., the daily 1-hour maximum background O_3 concentration) was factored into the calculation of the rollback factor.

The modeling efforts cited by Staff estimated natural background O_3 concentrations within North America by removing all anthropogenic emissions of NO_x, CO, and nonmethane hydrocarbons (including NO_x emitted from aircraft and fertilizer, but not

biomass burning). Thus, the definition of background used in the models is different than the definition of *policy-relevant background* used by Staff (see Chapter 4). Unless the State of California plans to eliminate all anthropogenic emissions of NO_x, CO, and nonmethane hydrocarbons (including NO_x emitted from aircraft and fertilizer), the estimates for the range of *policy-relevant background* hourly average concentrations will be greater at most locations in California than those estimated by the models. The Staff, by subjectively selecting 0.04 ppm, has underestimated *policy-relevant background* at many areas across California. In addition, the use of a 4-hour afternoon average concentration is inappropriate for estimating the range of *policy-relevant background* O_3 concentrations for the daily maximum 1-hour period. Given the (1) different definitions used for background by CAAQSOD and the models, (2) low spatial resolution of the modeling results in large uncertainties, and (3) multi-hour averaging (4-hour afternoon averages) of the hourly average concentrations, the hourly average concentrations of *policy-relevant background* near sea level and at higher elevations will be higher at many locations than estimated by Staff. The background modeling results are highly uncertain and the results are not applicable to the California standard-setting process.

S.3 Evaluating the Proportional Linear Rollback Model

Staff concluded that the rate of change in the concentrations *above background* was similar among the percentiles and that this observation justified its application of a constant percentage rollback to all sites within an air basin. In carefully reviewing the information in the Appendix for the rate of change in the concentrations *above background* from the 1980s, the data do not show a constant percentage change *above background* for each of the

percentiles for the sites provided in the Appendix. The pooled South Coast Air Basin results show the least variability of the percentage changes among the percentiles. However, it is not appropriate to pool the information because this smoothes the variability and provides an optimistic picture of the observation. It is important to inspect the variability across the percentiles site by site and then, based on the results, draw conclusions as to whether the application of a constant percentage rollback to all sites within an air basin is justified. Based on the data presented in the Appendix of Chapter 10, it appears that the variability is too great to assume that a constant percentage rollback is justified for the South Coast Air Basin.

There appeared to be no attempt to justify the decision by Staff to use the same rollback model for all locations in California. Although trending data exist for the South Coast Air Basin, such data are not widely available in other geographic areas in California. The greatest progress in the reduction of the peak hourly average concentrations has been made in the South Coast Air Basin. The fact that there has been no attempt to justify the use of the rollback model across all of California means that there is no evidence that even if the rollback model were applicable in the South Coast Air Basin, it would be useful outside that geographic area. In addition, the rollback model used in the health benefits analysis does not necessarily mimic the slowing down of the mid-level hourly average concentrations in relation to the higher hourly average values and therefore, the predicted distribution of the concentrations may not be reliable.

S.4 The Use of Epidemiologic Data in the Benefits Analysis

Staff has relied on epidemiological results to estimate health benefits. The available epidemiological evidence on O_3 is highly uncertain. The inconsistencies and model

dependency of effect estimates, although recognized by Staff, are many times ignored. Switzer (2004) noted that without a clear understanding of the reasons for inconsistent effects estimates, one cannot rule out the possibility that O_3 effect estimates are model artifacts. As such, current epidemiological studies have not established cause-effect relationships and are therefore not suitable for estimating potential benefits related to O_3 control.

Based on a review of the literature, Switzer (2004) reached the following conclusions:

1. **Sensitivity of O_3 effect estimates to model specification.** This issue was brought to light in the HEI reanalysis in the context of time and weather adjustments, and serves as a cautionary tale. The reported effects of O_3 are often difficult to discern and are inconsistent among cities, regions, seasons, and time lags. Such inconsistencies may be suggestive of modeling inadequacies, particularly in regard to unmodeled confounding and unexplained effect modifiers.

   That O_3 effect estimates are delicate is not surprising given that they are superimposed on much stronger effects (e.g., concomitant weather variations). Without a clear understanding of the reasons for inconsistent effects estimates, we cannot rule out the possibility that O_3 effect estimates are model artifacts.

2. **Enforced additivity in the analysis model**. The analysis models relied on by CAAQSOD assume that O_3 effects are necessarily the same at any temperature, even when restricted to summer data. Approaches to mitigate the problem, depending on availability of data, include joint response surface modeling of O_3 and its confounders or stratification of the analyses based on confounder categories.

3. **Enforced linearity of exposure-response**. Because O_3 health effect estimates are inconsistent across studies, cities, seasons, etc., putative benefits of ambient O_3 mitigation are difficult to know. Enforced model linearity of exposure-response, as in the case of the analysis models the CAAQSOD relies on, conceals heterogeneity of response. Pooling of response functions to obtain linearity is not statistically justified and leads to regulatory dilemmas.

4. **Spatial variability of O_3 health effect estimates within cities.** There has been insufficient attention to the issue of spatial variability of effect estimates within cities based on selection or combination of monitors.

5. **Incomplete characterization of the relations between ambient O_3 exposure, individual PM exposure, individual PM susceptibility to health effects, and community level health effect measures**. The models that CAAQSOD uses for the analysis of community health effects of O_3 do not have any link to individual response functions.

6. **Unresolved inconsistencies of O_3 effect estimates.** The following inconsistencies are unresolved: seasonal differences, regional grouping, spatial heterogeneity both between cities and within cities, time lag selection, and treatment of gaseous pollutant confounders.

Ultimately, the question that must be answered is this: If the time-series data are more important than the clinical results for establishing benefits estimates, are the data good enough to use in the decision-making process? Based on the time-series evidence presented, one simply cannot draw comfortable conclusions regarding the circumstances and magnitudes of ambient O_3 health effects, or whether reported O_3 health effects are causative. Thus, one might conclude that there is still too much uncertainty remaining in the epidemiological time-series results. In addition, many of the concerns expressed by Staff about the strengths and limitations of the extensive body of epidemiologic evidence of associations between health effects and air pollutants have not been adequately addressed. The growing pattern of inconsistent and inconclusive findings using time-series data is troublesome and presents policymakers with a very difficult decision about setting policy based on study results of questionable scientific validity.

S.5 Accruing Health Benefits Below the Proposed Standards

An important issue that was not addressed is what is the scientific rationale for accruing health benefits below the level of the proposed standards? The proposed California standards are based on primarily on chamber studies. The chamber studies do not show statistically significant effects below the 8-hour, 0.08 ppm level, which led Staff to propose a standard of 0.07 ppm, including a margin of safety. No rationale is provided for accumulating health benefits below the level of the standard to an assumed *policy-relevant*

background level of 0.04 ppm. To accrue health benefits below the level of the proposed standards implies acceptance of epidemiological data as demonstrating and quantifying causal relationships, when in fact, large uncertainties have been documented by Staff. Staff's comments preclude using current epidemiology data for this purpose. Staff has indicated that if there were no effects below the proposed O_3 standard, the health benefits analysis would be overestimated (page 10-13).

S.6 Conclusions

Based on the information provided in Chapter 10, the conclusion is that there are serious deficiencies in the health benefits analysis because (1) selecting a *policy-relevant background* 1- and 8-hour average concentration of 0.04 ppm for all geographic areas in California has underestimated the value, (2) selecting a proportional linear rollback model for estimating the distribution of the concentrations that result from attaining the 1- and 8-hour proposed standards will produce unreliable estimates, (3) using a highly uncertain methodology in the epidemiological studies will introduce large uncertainties in the accrued health benefits, and (4) accruing health benefits at O_3 concentrations between the proposed standards and background will provide large overestimates of benefits.

The health benefits estimates provided by Staff are highly uncertain and therefore, any policy decisions based on the benefits analysis should be considered speculative. Simply stated, current epidemiological data do not provide an adequate foundation upon which one can base a health benefits analysis. The current version of the benefits analysis simply represents a modeling exercise that may not provide realistic benefit estimates.

1. Introduction

The staff of the Air Resources Board (ARB) and Office of Environmental Health Hazard Assessment (OEHHA) has attempted to quantify in Chapter 10 of the California Ambient Air Quality Standard for Ozone Document (CAAQSOD) the adverse health effects of current O_3 levels in California by estimating the health benefits that would accrue from a hypothetical control strategy that achieves the proposed ambient air quality standards for O_3. The Staff has pointed out that estimating the health benefits associated with reductions in levels of ambient O_3 involves the following four elements:

1. Estimates of the changes in O_3 concentrations due to a hypothetical control strategy;
2. Estimates of the number of people exposed to O_3;
3. Baseline incidence of the adverse health outcomes associated with O_3.
4. Concentration-response (CR) functions that link changes in O_3 concentrations with changes in the incidence of adverse health effects. These functions produce a beta coefficient, indicating the percent reduction in a given health outcome due to a unit change in O_3.

The product of these elements generates estimates of the expected number of avoided adverse health outcomes associated with a hypothetical control strategy to reduce current levels of O_3 to the proposed California standard. The staff has focused on premature mortality, hospital admissions for respiratory diseases, emergency-room visits for asthma, school absences, and minor restricted activity days.

Most health studies considered in Staff's analysis were conducted with O_3 levels measured as 1-hour maximum or 8-hour maximum. However, there were some studies that measured O_3 averaged over other time increments. Because these studies were conducted

throughout the United States and other parts of the world, Staff used a national average of adjustment factors to convert all measurements to 1-hour and 8-hour averages. Staff assumed that the 1-hour maximum was 2.5 times the 24-hour average, and 1.33 times the 8-hour average concentration.

The proposed California standards are based primarily on chamber studies. These studies have been given primary focus because both the dose and response are well characterized. Staff noted in Chapter 8 (page 8-20) that although epidemiological associations have been reported for outcomes including cardiovascular mortality and hospital visits for children less than age two, it is difficult to attribute these adverse outcomes to a specific O_3 concentration or time. Staff notes that because of the high temporal correlation of 1-, 8-, and 24-hour average O_3, the averaging time of concern cannot be discerned from these studies. In addition, most of the studies used linear (i.e., no-threshold) models and did not explicitly test for thresholds. As noted by Staff, certain models, such as the time-series studies of mortality and hospitalization, suffer from problems of confounding from seasonal and weather factors and possibly co-pollutants. Thus, although epidemiological data were available, Staff chose to focus on the chamber studies for proposing the level of the two O_3 standards.

For performing the health benefits analysis in Chapter 10, Staff has relied on epidemiological results, even though Staff notes that are a variety of unresolved statistical issues associated with the use of epidemiological results. Several challenges and unresolved issues present themselves with respect to designing and interpreting time-series studies of O_3-related health effects. As noted by Staff in Section 10.6 (beginning on page 10-11), the principal challenge facing the analyst in the daily time-series context is to remove bias due to

confounding by short-term temporal factors operating over time scales from days to seasons. Staff notes that few studies to date have thoroughly investigated these potential effects with reference to O_3, introducing an element of uncertainty into the health benefits analysis.

Of particular importance is the strong seasonal cycle for O_3, high in summer and low in winter, which is opposite to the usual cycle in daily mortality and morbidity, which is high in winter and low in summer. Inadequate control for seasonal patterns in time-series analyses leads to biased effect estimates. Also, temporal cycles in daily hospital admissions or emergency room visits are often considerably more episodic and variable than is usually the case for daily mortality. As a result, smoothing functions that have been developed and tuned for analyses of daily mortality data may not work as well at removing cyclic patterns from morbidity analyses. Potential confounding by daily variations in co-pollutants and weather is another analytical issue highlighted by Staff. Another issue relates to the shape of the CR function and whether there is an effect threshold. As noted by Staff in its decision not to focus on the use of epidemiological data to propose the two standards, large uncertainties exist in the data.

For the purposes of its analysis, Staff has estimated health benefits down to an assumed background concentration of 0.04 ppm (except for emergency room visits for asthma, for which a higher threshold value was used). The selection of a *policy-relevant background* level of 0.04 is subjective and not supported by the information provided by Staff in Chapter 4. Approximately 76% - 86% of the benefits presented in Chapter 10 are accrued at O_3 concentrations between the proposed standards and background. Staff notes (page 10-13) that to the extent that there is a population threshold, the estimated benefits may

not be accurate. For reasons described in Section 3, the introduction of a non-linear model has important ramifications in the estimates associated with the modeling results.

A key issue that needs to be addressed is the scientific rationale for accruing health benefits below the level of the proposed standards. As noted above, the proposed California standards are based primarily on chamber studies. The chamber studies do not show statistically significant effects below the 8-hour, 0.08 ppm level. Thus, there is no rationale provided for accumulating health benefits down to an assumed *policy-relevant background* level of 0.04 ppm based on the clinical studies. To accrue health benefits below the level of the proposed standards implies the use of epidemiological data, which exhibits large uncertainties as documented by Staff. Staff has indicated that if there were no effects below the proposed O_3 standard, the health benefits analysis would be overestimated (page 10-13).

A thorough review was undertaken to evaluate the Staff's decision to (1) use 0.04 ppm as a the concentration for *policy-relevant background*, (2) apply a proportional linear rollback model, and (3) substitute unreliable epidemiological results for the chamber analyses used to propose the level of the two standards. Based on this review, there are serious deficiencies in the health benefits analysis because (1) selecting a *policy-relevant background* 1- and 8-hour average concentration of 0.04 ppm for all geographic areas in California has underestimated the value, (2) selecting a proportional linear rollback model for estimating the distribution of the concentrations that result from attaining the 1- and 8-hour proposed standards will result in unreliable estimates, (3) using a highly uncertain methodology in the epidemiological studies will introduce large uncertainties in the accrued health benefits, and (4) accruing health benefits at O_3 concentrations between the proposed standards and background will provide large overestimates of benefits. Based on the above

observations, the estimated health benefits are highly uncertain and therefore any policy decisions based on the benefits analysis should be considered speculative. The current version of the benefits analysis simply represents a modeling exercise that may not provide realistic estimates of risk.

2 *Policy-Relevant Background* and the Limitations of the Rollback Methodology

2.1 The Selection of a Relevant Policy-Relevant Background Concentration

One of the most important aspects of the O_3 health benefits analysis is the adequacy of the rollback methodology. If the methodology is not adequate, then the application of the rollback model will introduce a level of uncertainty that may make the health benefits analysis unreliable. For estimating the daily reductions in current O_3 concentrations that result at all monitoring sites due to a hypothetical control strategy, rollback factors from the 1-hour and 8-hour O_3 design values to the applicable standard were calculated for each air basin. The O_3 design value selected was the highest for the three-year period (2001 to 2003). A background O_3 concentration of 0.04 ppm (i.e., the daily 1-hour maximum background O_3 concentration) was factored into the calculation of the rollback factor. The rollback factor was assumed to apply to each site in the air basin for every day in a given year. The selection of the level of the *policy-relevant background* concentration affects the estimated health benefits. Staff has estimated effects down to a background concentration of 0.04 ppm for most of the various biological endpoints. As indicated in the Introduction, about 76 - 86% of the health benefits presented in Chapter 10 accrue at O_3 concentrations between the proposed standards and *policy-relevant background*. Thus, if one were only interested in the accrued

health benefits between the current levels and the proposed standards, the benefits would have been reduced from the estimated values to 14 – 24%.

Staff in Chapter 10 used an uncontrollable O_3 concentration of 0.04 ppm (the average daily 1-hour maximum) and factored this concentration into the calculation of the rollback factor. In Chapter 4, the CAAQSOD states that from a regulatory perspective, the important distinction is not between "natural" and "anthropogenic" O_3, but between O_3 produced by controllable emissions and O_3 due to emissions beyond the reach of regulation. Anthropogenic O_3 produced outside the jurisdiction of an agency and transported into a control region is functionally indistinguishable from that due to natural processes. Within the range of concentrations due to such external or uncontrollable sources, those concentrations that may impact determinations of compliance with air quality standards or limit the potential air quality improvements due to control programs have been defined by the CAAQSOD as *policy-relevant background*. The selection of an appropriate level for the *policy-relevant background* is critical to the health benefits analysis. It is important to stress that Staff has used 0.04 ppm as the *policy-relevant background* O_3 concentration. The subjectively determined value may be too low based on the discussion in Chapter 4.

Estimates of *policy-relevant background* concentrations need to consider the important contribution from stratospheric O_3, as well as other natural sources. There is a large variability among global models on the attribution of the contribution of natural O_3 to the background. On page 4-11, Staff states that it appears that "background" O_3 in California is dominated by natural tropospheric and stratospheric processes. Staff (page 4-11) concluded that the average **natural background** O_3 *near sea level* is in the range of 15-35 ppb (4-hour afternoon average concentration), with a maximum of about 40 ppb (4-hour afternoon

average). The range of concentrations is based on the Fiore *et al*. (2002) modeling exercise. Staff believed that exogenous enhancements to "natural" levels were generally small (about 5 ppb). Thus, Staff concluded that **natural background** concentrations (not *policy-relevant background* concentrations) were in the 40-45 ppb range (4-hour afternoon averages).

The Fiore *et al*. (2002, 2003) modeling efforts estimated natural background O_3 concentrations within North America by removing all anthropogenic emissions of NO_x, CO, and nonmethane hydrocarbons (including NO_x emitted from aircraft and fertilizer, but not biomass burning). Thus, the definition of background used by Fiore *et al*. (2002, 2003) is different than the definition of *policy-relevant background* used by Staff. Unless the State of California plans to eliminate all anthropogenic emissions of NO_x, CO, and nonmethane hydrocarbons (including NO_x emitted from aircraft and fertilizer), the estimates for the range of *policy-relevant background* hourly average concentrations will be greater at most locations in California than those estimated by Fiore *et al*. (2002, 2003). Thus, Staff, by subjectively selecting 0.04 ppm, has underestimated *policy-relevant background* at many areas across California. In addition, the use of a 4-hour afternoon average concentration is inappropriate for estimating the range of *policy-relevant background* O_3 concentrations for the daily maximum 1-hour period. Given the (1) different definitions used for background by CAAQSOD and Fiore *et al*. (2002, 2003), (2) low spatial resolution of the Fiore model that results in large uncertainties, and (3) multi-hour averaging (4-hour afternoon averages) of the hourly average concentrations, the hourly average concentrations of *policy-relevant background* near sea level and at higher elevations will be higher than stated by Staff at many locations. Recognizing that Fiore *et al*. (2003) noted that natural background levels never exceed 0.04 ppm, and given empirical observations during springtime in which the influence

of stratospheric O_3 on surface exposures at various locations in the world has been documented, the modeling results published by Fiore *et al*. (2002, 2003) should be considered highly uncertain and the results not applicable to the California standard-setting process.

In reviewing the summary of the health benefits analysis, over 80% of the benefits for each of the selected endpoints are accrued in the South Coast, San Joaquin Valley, Mojave Desert, and Sacramento Valley Air Basins. Three of the four basins are inland. There are biogenic emissions in the San Joaquin Valley and Sacramento Valley and surrounding foothills that may be responsible for enhancing the *policy-relevant background* O_3 levels. Modeling has indicated that removal of anthropogenic emissions will not necessarily reduce O_3 levels below 0.06 ppm on days when historic O_3 levels were high for the 1-hour standard. For the San Joaquin Valley, the major population centers (and suburbs around them) are Bakersfield, Fresno, Merced, Modesto, plus lesser towns along Route 99. Aside from biogenics, soil NO_x emissions are associated with agricultural sources. Its emissions are related to soil moisture and are cyclical with irrigation. Thus, looking at irrigation cycles and meteorology, it is highly likely that on some days there is soil NO_x, when O_3 formation is conducive. Both dairy and feedlots exist in the area, which may also contribute to the NO_x emissions and *policy-relevant background* levels.

As part of the California Regional Air Quality Studies, the San Joaquin Valley Air Quality Study was carried out in 1990 with collection of a robust aerometric data designed for model evaluation and use. Numerous simulations using the August 2-6, 1990 episode were made over time, with continued improvement in model inputs and performance during this period. A set of sensitivity simulations was made in which the domain-wide

anthropogenic emissions were zeroed out. Even under these very extreme conditions, exceedances of the Federal 8-hour 80 ppb standard occurred in some locations within the domain, noticeable within the Mountain Counties.

P*olicy-relevant background* concentrations may vary during the spring, summer, and fall. Thus, these concentrations will have to be characterized so that the rollback model takes into consideration the season-specific *policy-relevant background* concentrations. The application of a 0.04 ppm background level across the State of California by staff is too low for use in the rollback models. From a regulatory perspective, it is important to identify the variability of O_3 hourly average concentrations associated emissions that are beyond the reach of regulation. Both the maximum hourly average and the 8-hour average will be influenced by the variability of hourly *policy-relevant background* O_3 concentrations. The empirical data provide a solid indication to CAAQSOD that *policy-relevant background* O_3, as defined on page 4-1, will be higher than 40 ppb in many locations in California during specific times of the year.

2.2 The Limitations of the Rollback Methodology

2.2.1 The Peaks Are Coming Down Faster Than the Mid-Level Concentrations

The rollback methodology used by Staff assumed that under the hypothetical attainment setting, all O_3 observations within an air basin were subjected to the same percentage rollback factor based on the basin's three-year high value. As indicated earlier, background was defined as 0.04 ppm for both the 1-hour daily maximum and the 8-hour daily maximum concentrations (see equations on page 10-23). To investigate the plausibility of this assumption, staff examined the trends in the annual distributions of the 1-hour and 8-

hour concentrations of O_3 in the South Coast Air Basin (SoCAB). Due to its population and current O_3 levels, a significant proportion of statewide health benefits are projected to accrue in the SoCAB. The maximum, the 90th, 80th, 70th, 60th, 50th and 40th percentiles from the annual distribution of the basin's daily high concentrations as well as the individual site's daily highs, according to Staff, showed a downward trend from the 1980s (Figures 2-1 and 2-2). When Staff examined the rate of change in the concentrations above background from the 1980s, it concluded that the rate of change above background was similar among the percentiles. Staff believed that this observation justified its application of a constant percentage rollback to all sites within an air basin (see page 10-3). Results for several representative sites used in this analysis of O_3 trends can be found in the Appendix in Chapter 10.

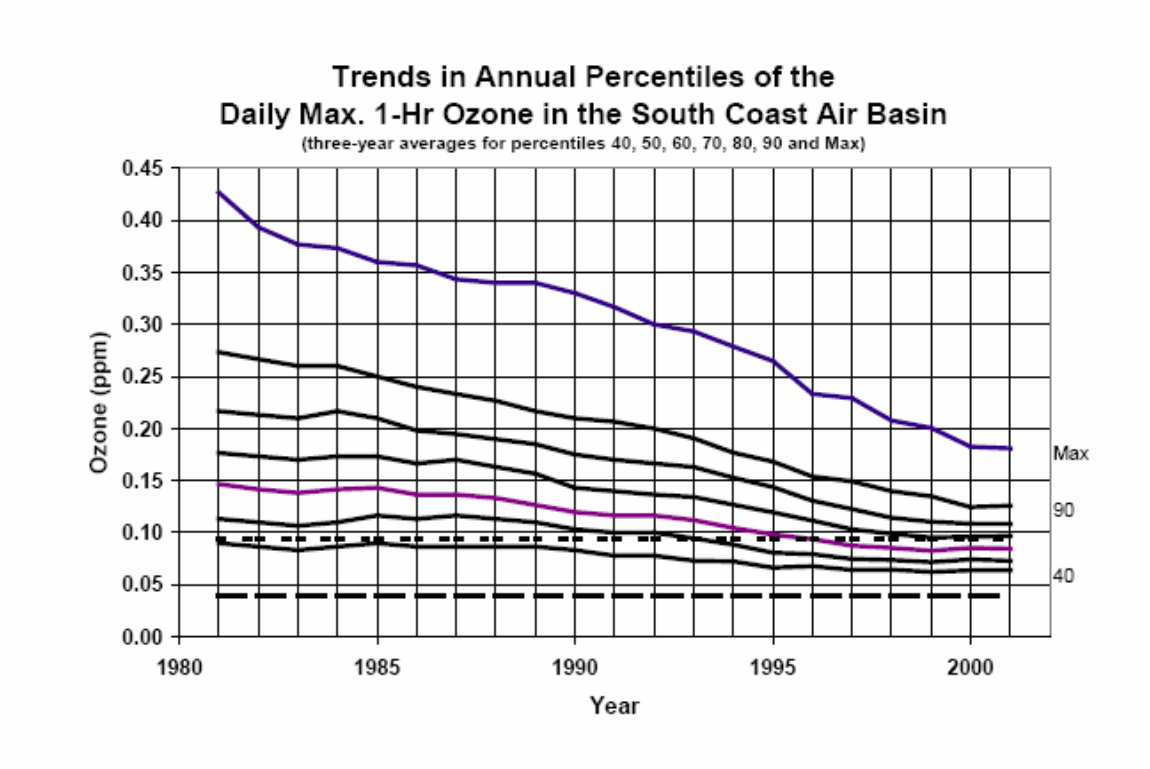


Figure 2-1 Trends in Annual Percentiles of Daily Max 1-hour Ozone in the South Coast Air Basin (page 10-25). Figure 10-1 in Appendix).

Note that in Figures 2-1 and 2-2, the slopes of the percentile lines are not parallel to one another. As one studies the slopes of the maximum, the 90th, 80th, 70th, 60th, 50th and 40th

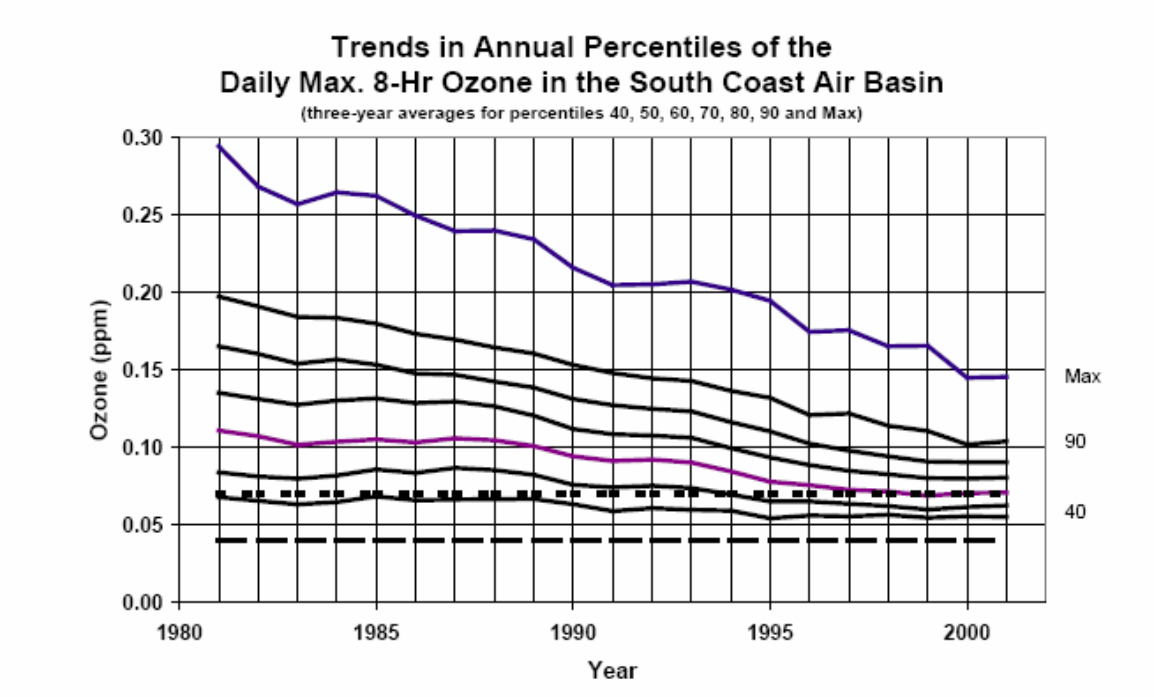


Figure 2-2 Trends in Annual Percentiles of Daily Max 8-hour Ozone in the South Coast Air Basin (page 10-27). (Figure 10-2 in Appendix)

percentiles, one finds that the general tendency is for the slopes to become more horizontal (i.e., parallel to the x-axis) as the percentiles go from high to low. Staff characterized the rate of change in the concentrations *above background* for the South Coast Air Basin from the 1980s and concluded that they were similar among the percentiles (page 10-3). Whether Staff was correct in assuming that the rate of change *above background* was similar across the percentiles will be discussed shortly. First, we will discuss the ramifications associated with the tendency for the trending percentile slopes to become more horizontal as the concentrations decrease.

The atmosphere's response to changes in O_3 precursors is a *non-linear* process. This means that the hourly concentrations within an 8-hour average do not respond in a linear manner when emission reductions occur. The changing slopes observed in Figure 2-1 (daily maximum 1-hour concentrations) and Figure 2-2 (daily maximum 8-hour concentrations) illustrates the disproportionate reduction in the hourly average concentrations as a function of absolute value. Evidence for the disproportionate reduction in hourly average concentrations of O_3 as emission reductions occur can be observed in Section 10.9 (Appendix). Figure 2-3 (Figure 10-3 in the Appendix) illustrates the trends in the annual percentiles of the daily maximum 1-hour O_3 concentrations at the North Long Beach monitoring site. For those concentrations below 0.09 ppm (the 40^{th} – 90^{th} percentiles), the trending lines are close to horizontal, which implies very little reduction in values, even though the concentrations above 0.10 ppm were reduced. Similar results occurred when the daily maximum 8-hour figure was reviewed.

Similar results are illustrated in Figure 2-4 (Figure 10-5 in the Appendix) for the L.A. North Main Street monitor. The trends in the annual percentiles of the daily maximum 1-hour O_3 concentrations at the monitoring site show that the concentrations below 0.09 ppm are showing little change over the 1980 – 2001 period, even though substantial reductions were occurring to the concentrations above 0.10 ppm. Figure 2-5 (Azusa) (Figure 10-7 in the Appendix) and Figure 2-6 (Crestline) (Figure 10-9 in the Appendix) show the same pattern of fairly large reductions in the hourly average concentrations above 0.10 ppm, but little change to those concentrations in the range below 0.10 ppm.

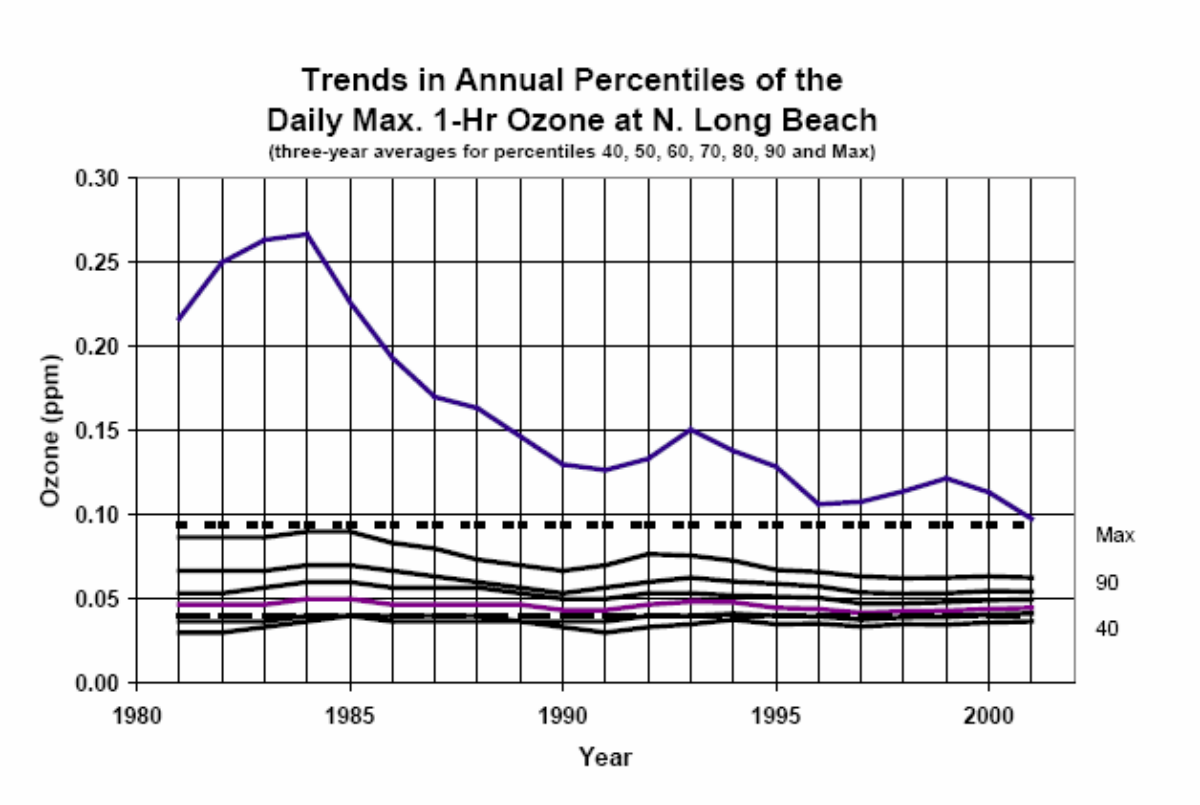


Figure 2-3 Trends in Annual Percentiles of Daily Max 1-hour Ozone at N. Long Beach. Figure 10-3 in Appendix.

Lefohn *et al*. (1998) discussed the disproportionate rate of reduction of the hourly average concentrations. Using the EPA's Aerometric Information Retrieval System (AIRS), Lefohn *et al*. (1998) reported that for the period 1993-1995, approximately 50% of the areas that violated the 8-hour standard were influenced by 4 or more occurrences of mid-level hourly average concentrations (i.e., 60 - 90 ppb). In addition, the authors identified those sites that demonstrated a significant reduction in O_3 levels for the period 1980-1995. Using the data from the sites that experienced reduced O_3 levels over the period of time, Lefohn *et al*. (1998) investigated whether the rate of reduction of the mid-level hourly average concentrations was similar to the rate experienced by the high hourly average concentrations.

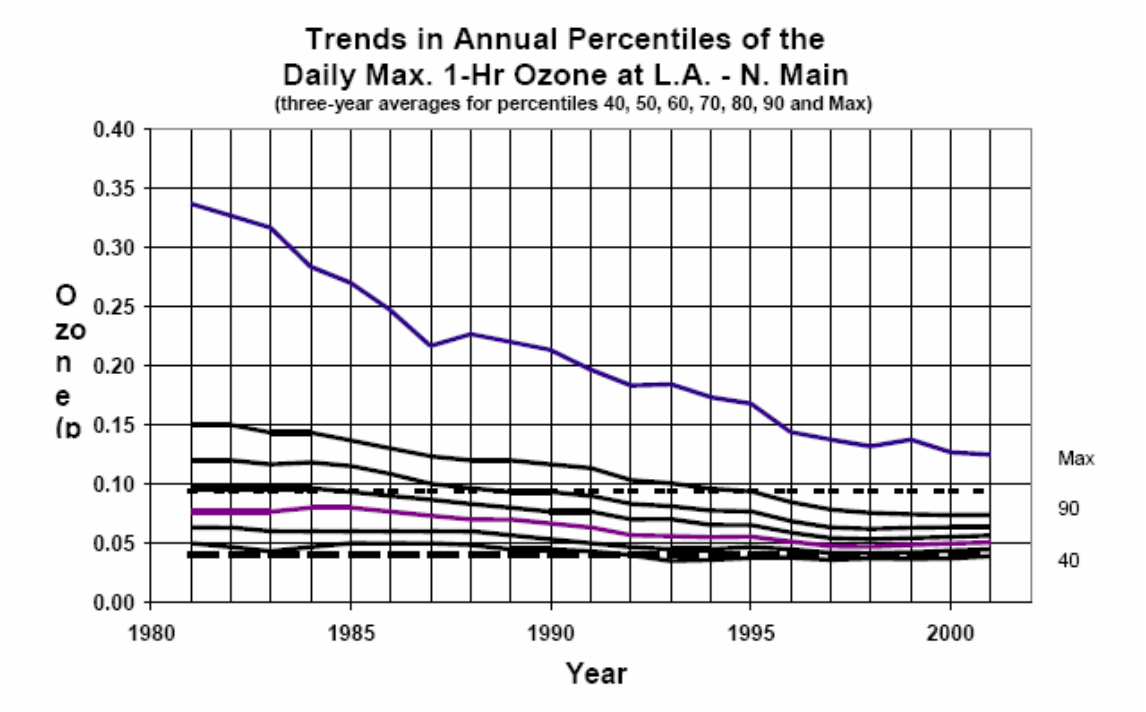


Figure 2-4 Trends in annual percentiles of daily max 1-hour ozone L.A. – N. Main (Figure 10-5 in Appendix).

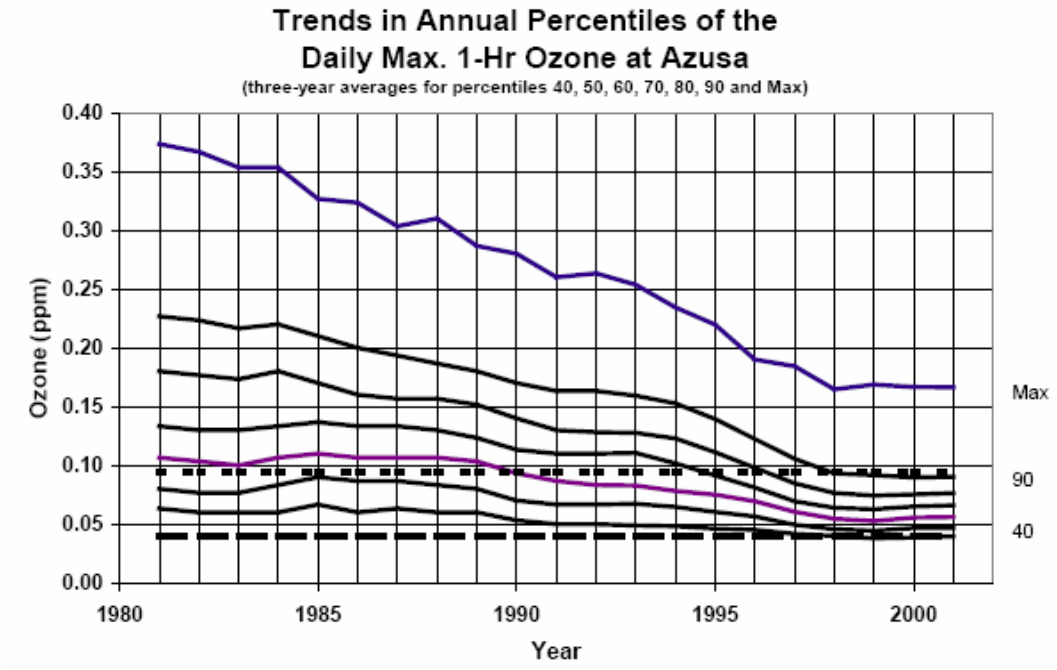


Figure 2-5 Trends in annual percentiles of daily max 1-hour ozone at Azusa (Figure 10-7 in Appendix).

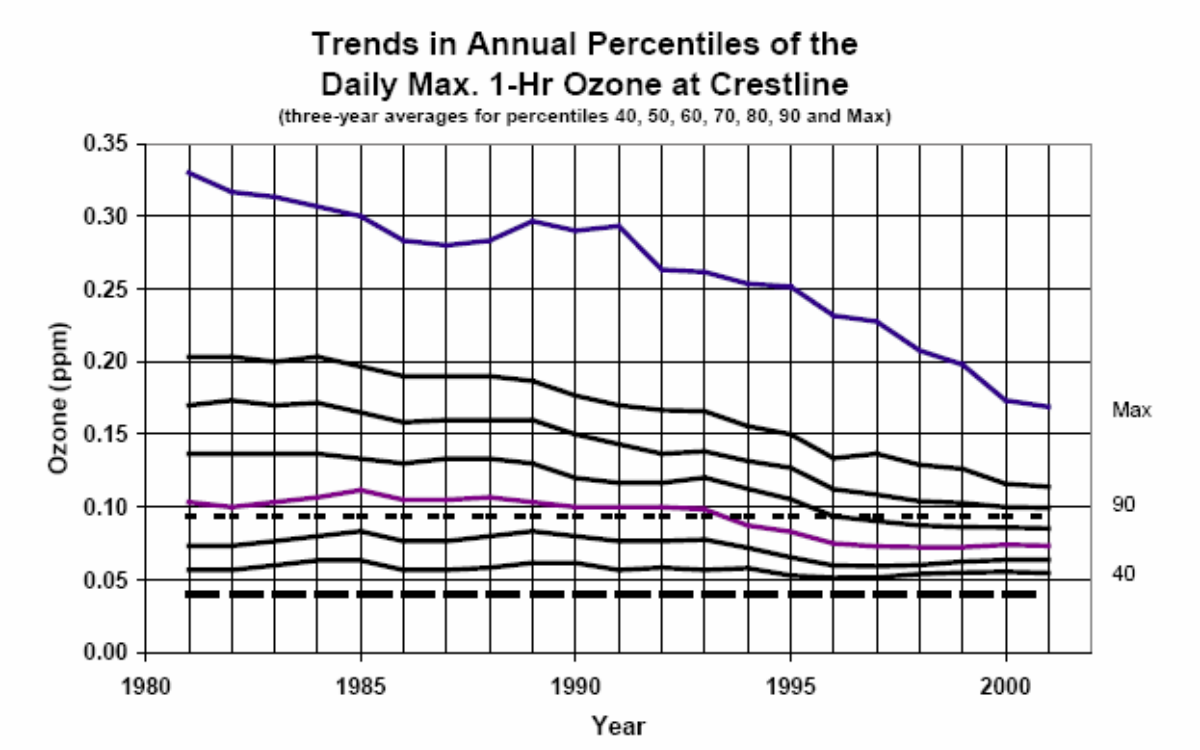


Figure 2-6 Trends in annual percentiles of daily max 1-hour ozone at Crestline (Figure 10-9 in Appendix).

The analysis indicated that the hourly average concentrations in the mid range were reduced more slowly than the hourly average concentrations above 90 ppb. Figure 2-7 is an example that shows that the higher hourly average concentrations (i.e., above 90 ppb) decreased at a faster rate (greater negative rate per year) than the hourly average concentrations in the mid-level range. The numbers of hourly average concentrations in the low end of the distribution also decreased. Apparently, both the high and low ends of the distribution were moving toward the center of the distribution.

As control strategies are implemented, the rate of change to reduce the higher hourly average concentrations will be greater than the rate of change of the mid-level concentrations. As mentioned earlier, this observation is illustrated in Figure 10-1 and 10-2 shown in the CAAQSOD. Figure 2-7 from Lefohn *et al*. (1998) illustrates the

disproportionate reduction of the hourly average concentrations. Note that for a monitoring site in Ventura County, actual reduction in emissions resulted in both the frequency of the higher hourly and lower hourly average concentrations being reduced. The frequency of occurrence of the mid-level hourly average concentrations (i.e., the 50 – 70 ppb range) increased. This meant that both ends of the hourly average concentration distribution were "squeezed" towards the middle of the distribution. The result is that as serious emission reductions occur, at some monitoring sites in California there will be an increase in the mid-level hourly average concentrations. Lefohn *et al*. (1998) identified a similar response in hourly average concentrations as a result of emission reduction for other California sites.

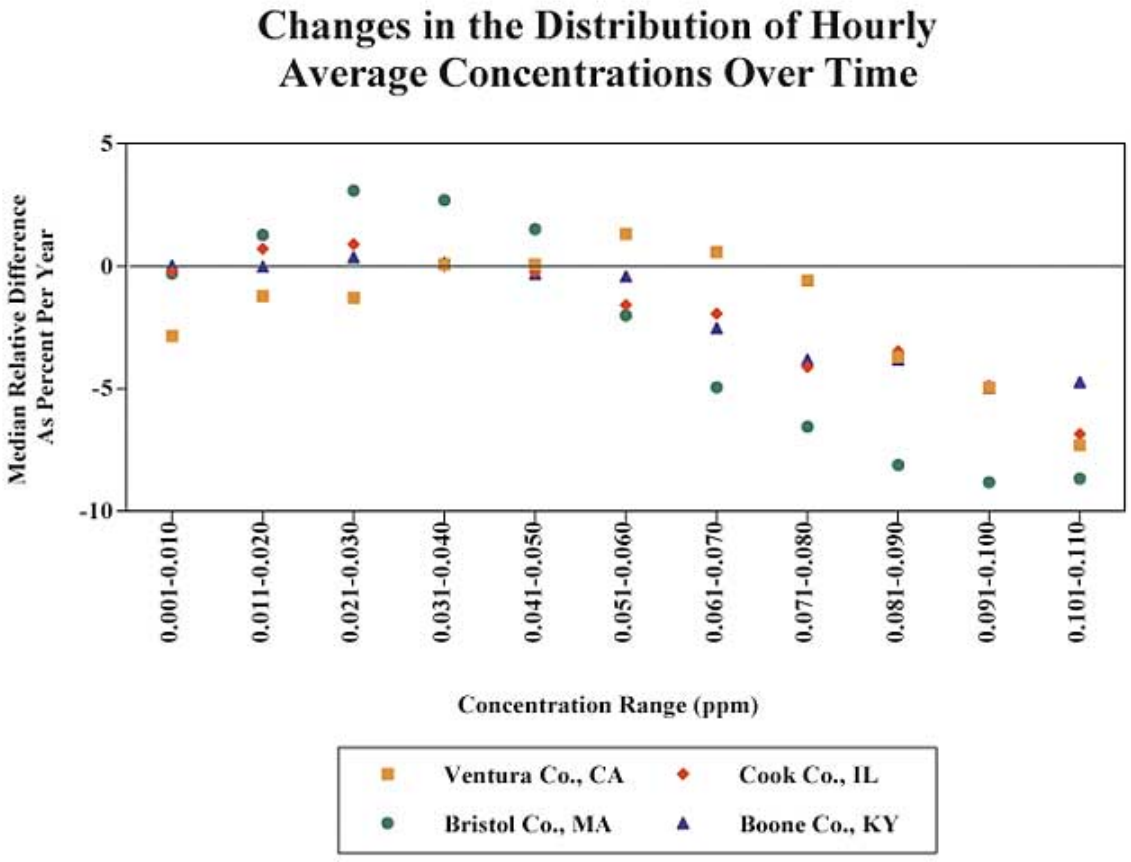


Figure 2-7. Changes in distribution of hourly average concentrations over time. Source: Lefohn *et al*. (1998).

Based on the results published by Lefohn *et al*. (1998), which used empirical data, it appears that when control strategies are implemented, the higher hourly average concentrations will be reduced faster than the mid-level values. Similar to the results obtained for the hourly values, the rate of decline for the 8-hour daily maximum values in the mid range will be much slower than the higher 8-hour values. This is observed in Figure 2-2. In addition, Figure 2-8 illustrates the slowing down process for Fairfield County, Connecticut. Note the rapid decrease in the early years and then a "flattening" of the curve in the later years.

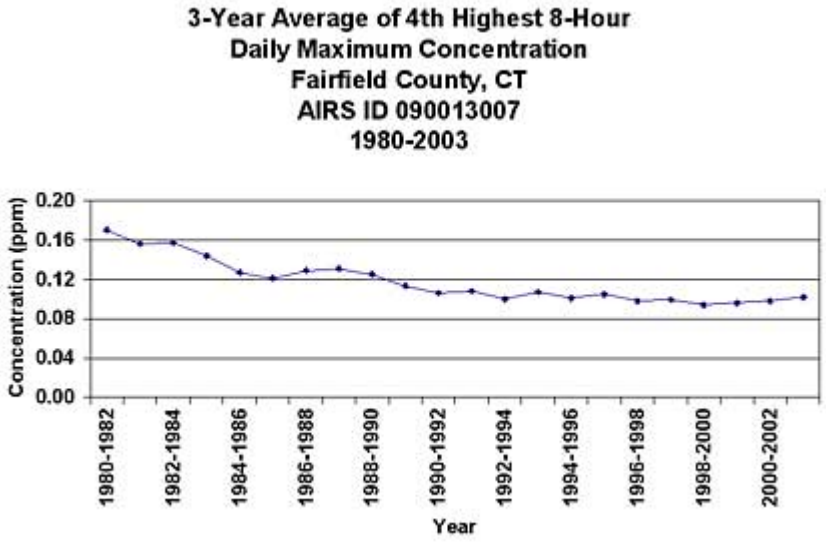


Figure 2-8. Trends of the 8-hour standard for the period 1980-2003.

This flattening has been observed nationwide, including California. The EPA has recently published the report, *The Ozone Report - Measuring Progress Through 2003* (EPA, 2003). The Agency notes that there has been a slowing down in the reduction of both the 1-hour and 8-hour average concentrations. Figures 2-9 and 2-10, taken from that report, illustrate the changes over the period 1980-2003.

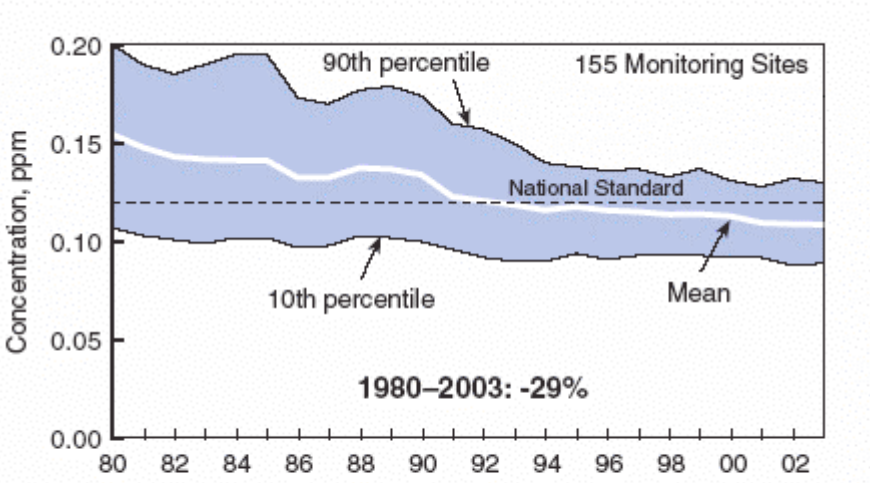


Figure 6. One-Hour Ozone Air Quality Trend, 1980–2003, Based on Running Fourth Highest Daily Maximum 1-Hour Ozone Value over 3 Years.

Figure 2-9. One-hour ozone air quality trend, 1980-2003, based on running 4th highest daily maximum 1-hour ozone value over 3 years. Source: *The Ozone Report - Measuring Progress Through 2003*.

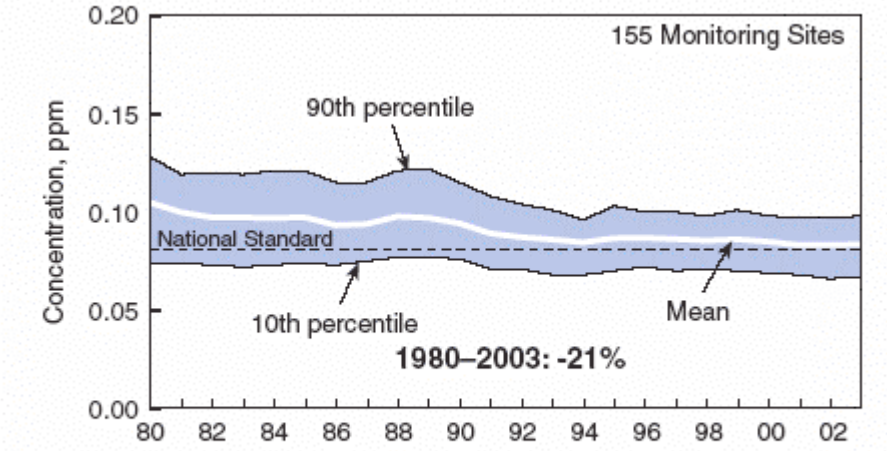


Figure 8. 8-Hour Ozone Air Quality Trend, 1980–2003, Based on 3-Year Rolling Averages of Annual Fourth Highest Daily Maximum 8-hour Ozone Concentrations.

Figure 2-10. 8-hour ozone air quality trend, 1980-2003, based on running 4th highest daily 3-year rolling averages of annual 4th highest daily maximum 8-hour ozone concentrations. Source: *The Ozone Report - Measuring Progress Through 2003*.

Although the rollback model used by Staff does not involve temporal considerations, the empirical data shown in the Appendix indicate a slowing down of the reductions in concentrations. Using models, several investigators have commented on the difficulty in reducing the mid- level hourly average concentrations, while reducing the fourth highest 8-hour average daily maximum concentration. Winner and Cass (2000) noted that the higher hourly average concentrations were reduced much faster than the mid-level values during simulation modeling for the Los Angeles area. Reynolds *et al*. (2003) analyzed ambient O_3 concentrations used in conjunction with the application of photochemical modeling to determine the technical feasibility of reducing hourly average concentrations in central California, using the 1990 August 3-6 San Joaquin Valley Air Quality Study episode.
The following four isopleths show how O_3 responds to precursor reduction. Each isopleth was created by using the modeling outputs for a specific combination of precursor reduction done in 10% increments (for example 10% VOC reduction and 20% NO_x reduction would be one of 121 possible combinations), followed by plotting the results. The axes represent the amount of reduction from the 1999 baseline emissions, with 100% of each precursor being in the top right-hand corner. Ozone concentrations are represented by the isopleths. These isopleths can be calculated for either the 1-hour average or the 8-hour average, and illustrate how O_3 would respond to reduction for the area specified. As can be seen in the Fresno isopleths (Figures 2-11 and 2-12), in order to achieve a 90 ppb O_3 concentration for the peak hour, 85% reduction of NO_x from the baseline is needed. However, for the same day, a 70 ppb concentration for the peak 8-hour requires about 92% reduction of NO_x.

Similarly, for the Bakersfield isopleths (Figures 2-13 and 2-14), the 90 ppb peak 1-hour O_3 concentration requires an 82% reduction of NO_x from the baseline. Yet, for the 70

ppb peak 8-hour O_3 concentration, a 86% reduction will be required. Similar findings were observed for other inland Central California locations (Reynolds *et. al.*, 2003), as well as major areas in the eastern half of the United States (Reynolds *et. al.*, 2004).

Reynolds *et al*. (2003) have commented on possible chemical explanations for the observation that more prominent trends in peak 1-hour O_3 levels occur than for trends in peak 8-hour O_3 concentrations or in occurrences of mid-level (i.e., 60 – 90 ppb) concentrations. The authors noted that when anthropogenic VOC and NO_x emissions are reduced significantly, the primary sources of O_3 precursors are biogenic emissions and CO from anthropogenic sources. Chemical process analysis results indicated that slowly reacting pollutants such as CO could be contributing on the order of 10 – 20% of the O_3 produced. Moreover, the authors noted that process analysis indicated that as NO_x was reduced, the process for O_3 formation became more efficient, producing more molecules of O_3 for each molecule of NO_x. That is to some extent, decreasing emissions were offsetting the increased effectiveness of making O_3.

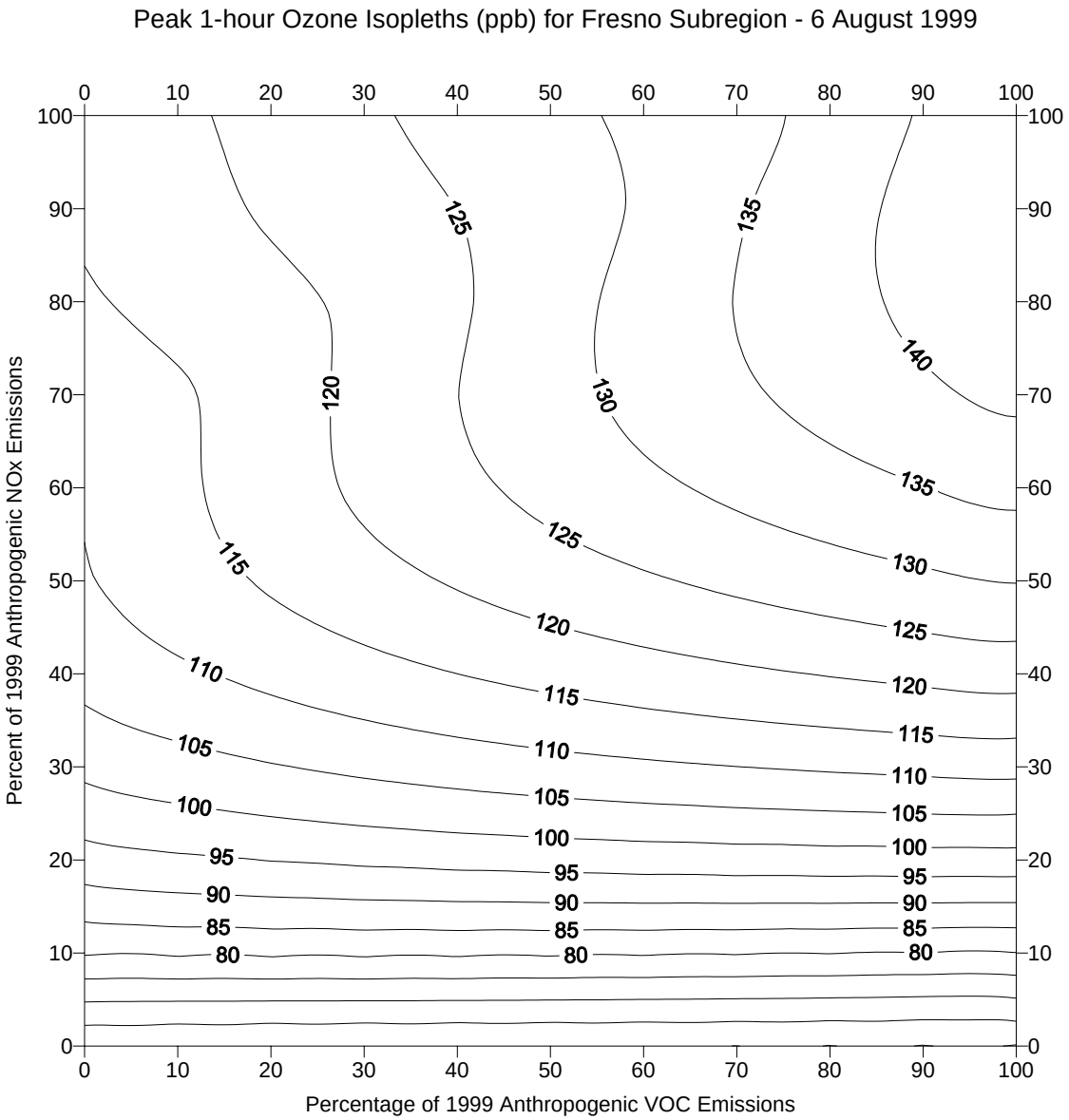


Figure 2-11. Peak 1-hour ozone isopleths (ppb) for Fresno Subregion August 6, 1999.

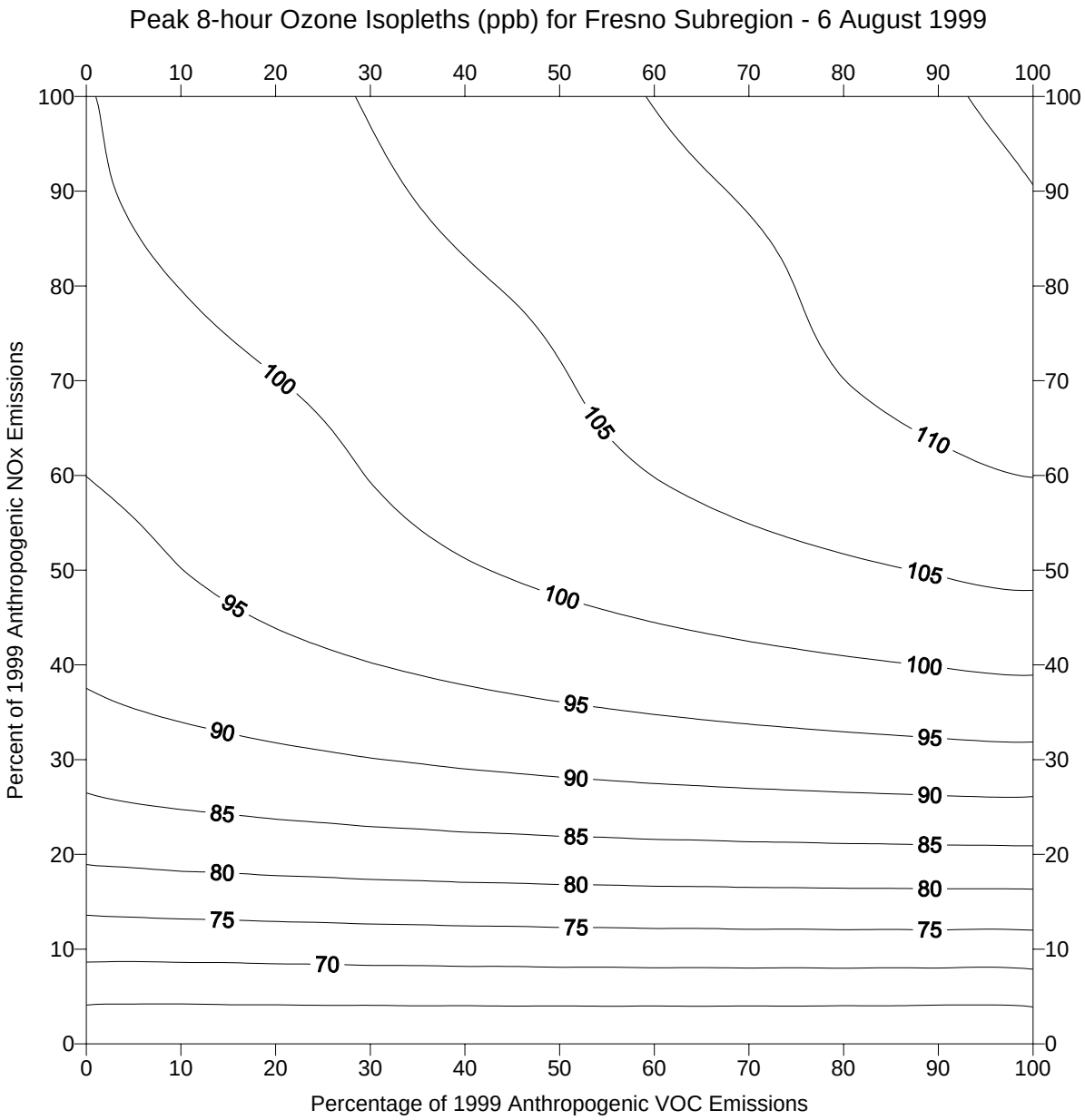


Figure 2-12. Peak 8-hour ozone isopleths (ppb) for Fresno Subregion August 6, 1999.

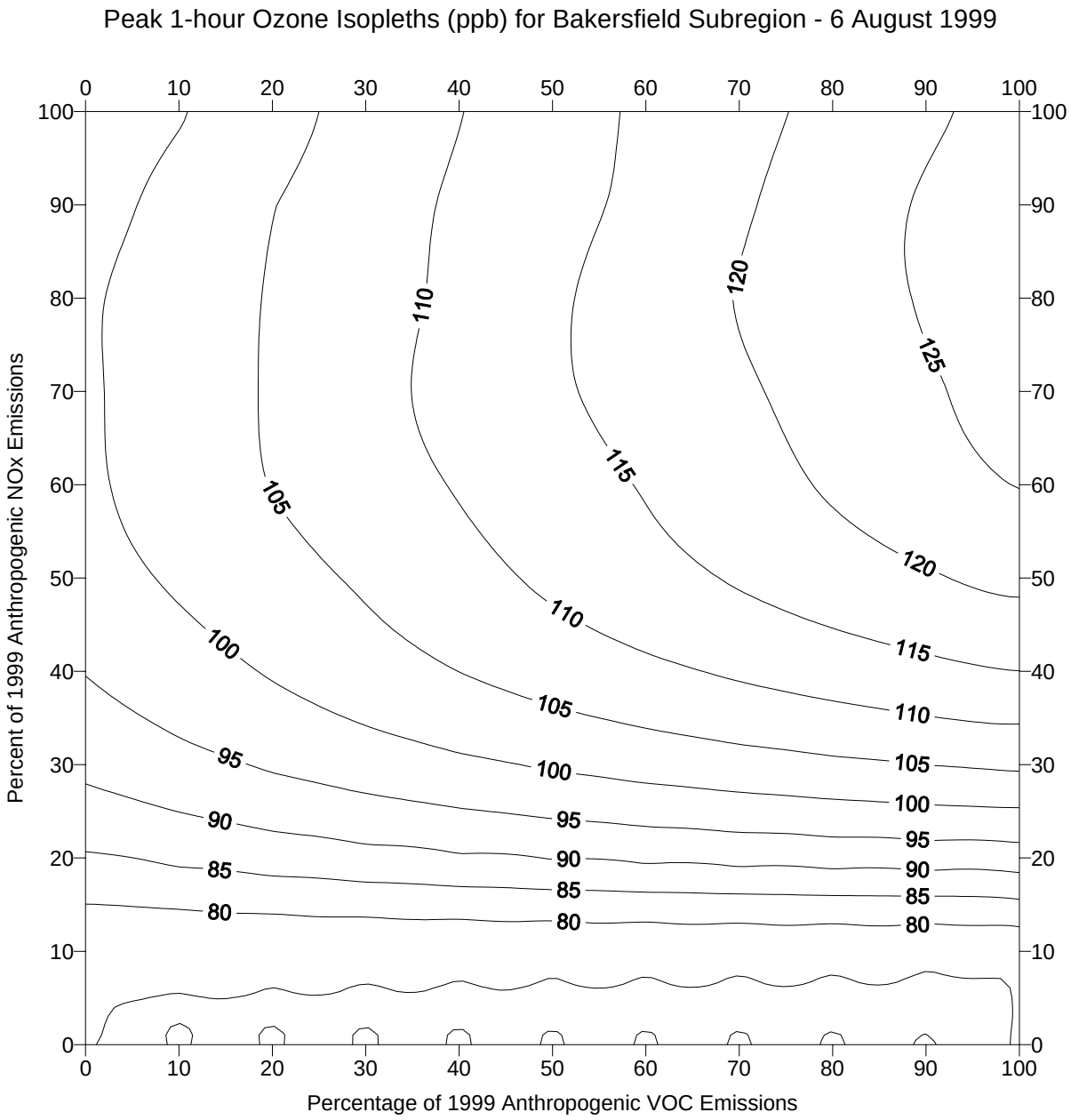


Figure 2-13. Peak 1-hour ozone isopleths (ppb) for Bakersfield Subregion August 6, 1999.

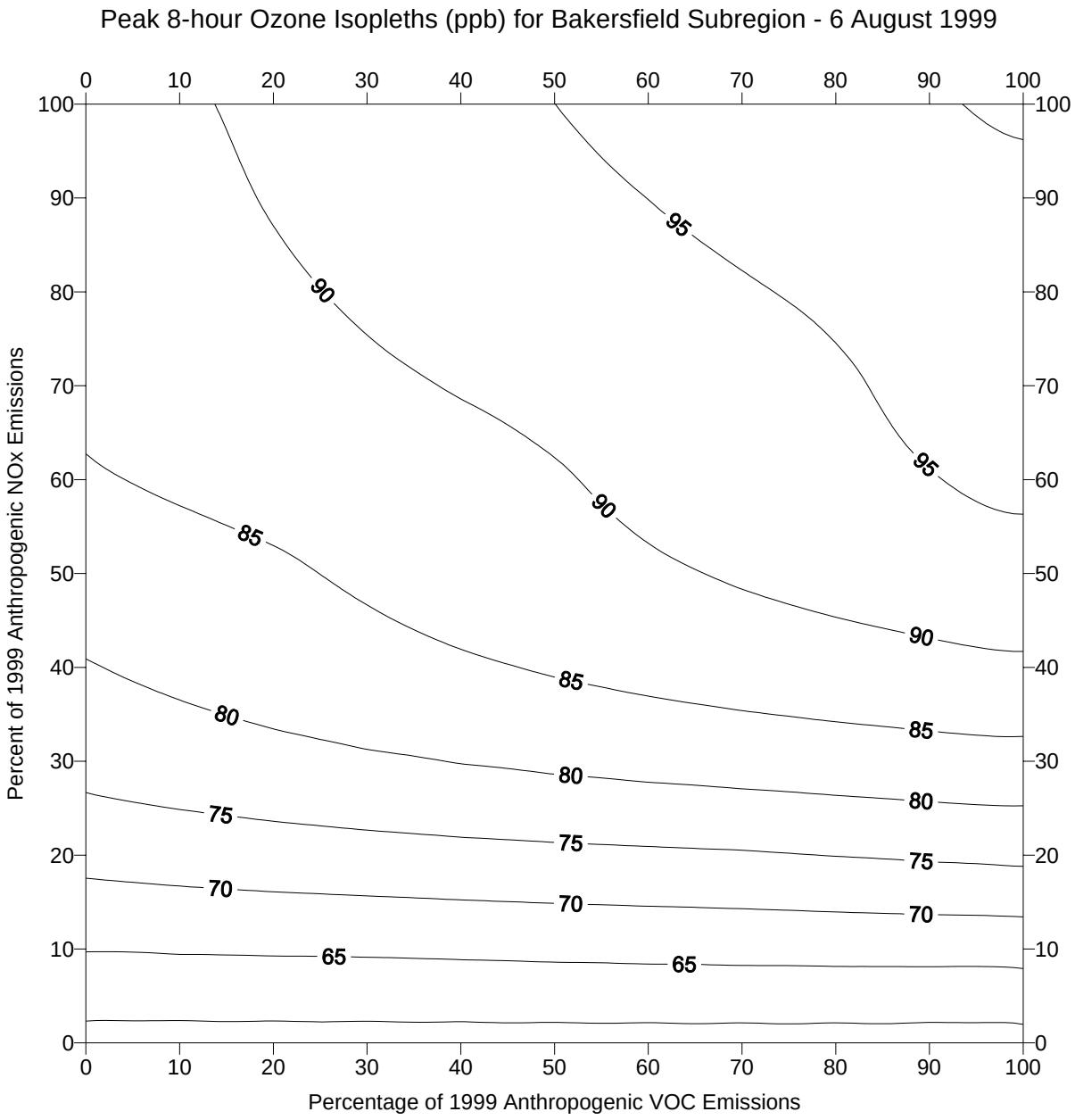


Figure 2-14. Peak 8-hour ozone isopleths (ppb) for Bakersfield Subregion August 6, 1999.

This was also seen in analysis of the ambient air quality data when looking at the change in the O_3 to NO_x ratio between 1990 and 2000, as shown in Figure 2-15. The ratio of O_3 to NO_x represents the efficiency of the atmosphere to produce O_3. There were significant NO_x emissions reductions in the Bakersfield area between 1990 and 1999, both for the

weekday (WD) and weekend (WE) days. Yet, with these large reductions, the ambient ratio shows an increase, and provides an independent correlation with the model results showing that the reduction of NO_x is offset by the efficiency of the atmosphere to produce more O_3 per molecule of NO_x, thus requiring more reduction of NO_x to meet the applicable O_3 standard. This was observed for other locations in Central California as well.

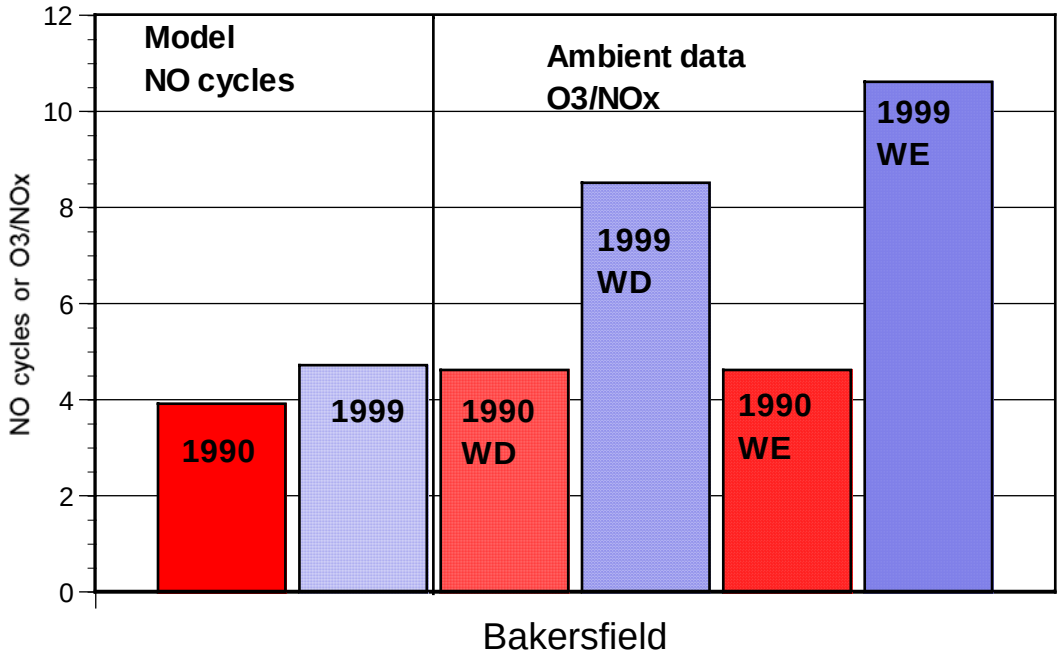


Figure 2-15. Change in the O_3 to NO_x ratio between 1990 and 2000.

Thus, the observation in the figures in the Appendix showing that the slopes of the low percentiles are becoming horizontal (the slowing down of the reduction of mid-level hourly average concentrations) has a chemical explanation. The rollback model used in the health benefits analysis does not necessarily mimic the slowing down of the mid-level hourly average concentrations.

2.2.2 The Rate of Change Above Background is Not Similar Among the Percentiles

As indicated earlier, Staff concluded that the rate of change in the concentrations *above background* was similar among the percentiles and this observation justified its application of a constant percentage rollback to all sites within an air basin. Table 2-1 (Table 10-6, page 10-26) shows an example of the summary of trends in annual percentiles of the daily maximum 1-hour O_3 in the South Coast Air Basin. Table 2-2 illustrates the percentage changes *above background* for the specific percentiles for the South Coast Air Basin. Table 2-3 (Table 10-10, page 10-34) shows an example of the summary of trends in annual percentiles of the daily maximum 1-hour O_3 for a site at North Main Street in Los Angeles. Table 2.4 illustrates the percentage changes above background for the specific percentiles for the monitoring site.

Table 10-6 Summary of Trends in Annual Percentiles of the Daily Max. 1-Hr Ozone in the South Coast Air Basin

	Average Value During Period		
Indicator	1980-1982	1990-1992	2000-2002
Maximum	0.427	0.317	0.183
Δ% above background		28%	63%
90th Percentile	0.273	0.207	0.125
Δ% above background		29%	64%
80th Percentile	0.217	0.170	0.109
Δ% above background		26%	61%
70th Percentile	0.177	0.140	0.096
Δ% above background		27%	59%
60th Percentile	0.147	0.117	0.086
Δ% above background		28%	57%
50th Percentile	0.113	0.100	0.075
Δ% above background		18%	53%
40th Percentile	0.090	0.078	0.064
Δ% above background		24%	52%

note: Delta % above background is the change in the portion of measured ozone since 1980-82 above "background", where background is defined as 0.04 ppm.

Table 2-1. Summary of trends in annual percentiles of the daily max. 1-h ozone in the South Coast Air Basin.

Table 2-2. The percentage change above background estimated from the rollback model for the South Coast Air Basin pooled data.

Indicator	**Δ Percent Above Background (1980-2002)**
Maximum	63%
90th Percentile	64%
80th Percentile	61%

70th Percentile	59%
60th Percentile	57%
50th Percentile	53%
40th Percentile	52%

Table 10-10 Summary of Trends in Annual Percentiles of the Daily Max 1-hour Ozone at L.A. - N. Main

	Average Value During Period		
Indicator	1980-1982	1990-1992	2000-2002
Maximum	0.337	0.197	0.127
Δ% above background		47%	71%
90th Percentile	0.150	0.113	0.074
Δ% above background		33%	69%
80th Percentile	0.120	0.090	0.064
Δ% above background		38%	70%
70th Percentile	0.097	0.077	0.055
Δ% above background		35%	73%
60th Percentile	0.077	0.063	0.050
Δ% above background		36%	74%
50th Percentile	0.063	0.050	0.044
Δ% above background		57%	84%
40th Percentile	0.050	0.043	0.037
Δ% above background		67%	100%

note: Delta % above background is the change in the portion of measured ozone since 1980-82 above "background", where background is defined as 0.04 ppm.

Table 2-3. Summary of trends in annual percentiles of the daily max. 1-h ozone at L.A. North Main.

Table 2-4. The percentage change above background estimated from the rollback model for the L.A. North Main data.

Indicator	Δ Percent Above Background (1980-2002)
Maximum	71%
90th Percentile	69%
80th Percentile	70%
70th Percentile	73%
60th Percentile	74%
50th Percentile	84%
40th Percentile	100%

Table 2-5 (Table 10-12, page 10-38) shows an example of the summary of trends in annual percentiles of the daily maximum 1-hour O_3 for a site at Azusa. Table 2-6 illustrates the percentage changes above background for the specific percentiles for the monitoring site. Table 2-7 (Table 10-14, page 10-42) shows an example of the summary of trends in annual percentiles of the daily maximum 1-hour O_3 for a site at Crestline. Table 2-8 illustrates the percentage changes above background for the specific percentiles for the monitoring site.

Table 10-12 Summary of Trends in Annual Percentiles of the Daily Max 1-hour Ozone at Azusa

Summary of Trends in Annual Percentiles of the Daily Max. 1-Hr Ozone at Azusa			
	Average Value During Period		
Indicator	1980-1982	1990-1992	2000-2002
Maximum	0.373	0.260	0.167
Δ% above background		34%	62%
90th Percentile	0.227	0.163	0.090
Δ% above background		34%	73%
80th Percentile	0.180	0.130	0.075
Δ% above background		36%	75%
70th Percentile	0.133	0.110	0.065
Δ% above background		25%	73%
60th Percentile	0.107	0.087	0.056
Δ% above background		30%	77%
50th Percentile	0.080	0.067	0.047
Δ% above background		33%	83%
40th Percentile	0.063	0.050	0.039
Δ% above background		57%	100%
note: Delta % above background is the change in the portion of measured ozone since 1980-82 above "background", where background is defined as 0.04 ppm.			

Table 2-5. Summary of trends in annual percentiles of the daily max. 1-h ozone at Azusa.

Table 2-6. The percentage change above background estimated from the rollback model for the Azusa data.

Indicator	Δ Percent Above Background (1980-2002)
Maximum	62%
90th Percentile	73%
80th Percentile	75%
70th Percentile	73%
60th Percentile	77%
50th Percentile	83%
40th Percentile	100%

Table 2-9 is a summary of the results from the Staff's analysis. In reviewing the information contained in the table, the data do not show a constant percentage change *above background* for each of the percentiles for sites within the South Coast Air Basin. Although Staff believed that the percent change above background (assumed 0.04 ppm level) was similar, it is not. The pooled South Coast Air Basin results show the least variability of the percentage changes among the percentiles. However, it is not appropriate to pool the information because this smoothes the variability and provides an optimistic picture of the observation. It is important to inspect the variability across the percentiles site by site and then, based on the results, draw conclusions whether the application of a constant percentage rollback to all sites within an air basin is justified. Based on the data presented in the Appendix of Chapter 10, it appears that the variability is too great to assume that a constant percentage rollback is justified for the South Coast Air Basin.

Table 10-14 Summary of Trends in Annual Percentiles of the Daily Max 1-hour Ozone at Crestline

	Average Value During Period		
Indicator	1980-1982	1990-1992	2000-2002
Maximum	0.330	0.293	0.173
Δ% above background		13%	54%
90th Percentile	0.203	0.170	0.116
Δ% above background		20%	53%
80th Percentile	0.170	0.143	0.100
Δ% above background		21%	54%
70th Percentile	0.137	0.117	0.086
Δ% above background		21%	52%
60th Percentile	0.103	0.100	0.074
Δ% above background		5%	46%
50th Percentile	0.073	0.077	0.064
Δ% above background		-10%	29%
40th Percentile	0.057	0.057	0.056
Δ% above background		0%	7%

note: Delta % above background is the change in the portion of measured ozone since 1980-82 above "background", where background is defined as 0.04 ppm.

Table 2-7. Summary of trends in annual percentiles of the daily max. 1-h ozone at Crestline.

Table 2-8. The percentage change above background estimated from the rollback model for the Crestline data.

Indicator	Δ Percent Above Background (1980-2002)
Maximum	54%
90th Percentile	53%
80th Percentile	54%
70th Percentile	52%
60th Percentile	46%
50th Percentile	29%
40th Percentile	7%

Table 2-9. A summary of the percentage change above background estimated from the rollback model for several monitoring locations listed in the Appendix.

Locations	Max	90th	80th	70th	60th	50th	40th
South Coast (pooled)	63	64	61	59	57	53	52
North Main	71	69	70	73	74	84	100
Azusa	62	73	75	73	77	83	100
Crestline	54	53	54	52	46	29	7

No attempt was made to justify the decision by Staff to use the same rollback model for all locations in California. Although trending data exist for the South Coast Air Basin, such data are not widely available in other geographic areas in California. The greatest progress in the reduction of the peak hourly average concentrations has been made in the South Coast Air Basin. The fact that there has been no attempt to justify the use of the rollback model across all of California means that there is no evidence that even if the rollback model were applicable in the South Coast Air Basin it would be useful outside that geographic area.

2.2.3 The Sensitivity of the Rollback Model to the Selection of a *Policy-Relevant Background* Concentration Level

It is important to understand how the rollback analysis was performed. To calculate changes in exposure to O_3 that reflect a hypothetical attainment of the proposed ambient air quality standards, a proportional linear rollback procedure was used. The magnitude of the rollback is determined by the following:

The rollback factor (RF) was calculated for each basin as follows:

If BasinMax > Std, then RF = (BasinMax - Std) / (BasinMax – BG).
If BasinMax <= Std, then RF = 0.

Then, for all sites within the basin, the portion of the site's current O_3 levels above background was adjusted by Staff as follows:

If OzCurrent > BG, then OzAttain = BG + (1 - RF) × (OzCurrent – BG). **(1)**
If OzCurrent <= BG, then OzAttain = OzCurrent.

Where:

OzCurrent = current daily O_3 observed value,
BasinMax = design value based on three years of measured data,
BG = background O_3 of 0.04 ppm,
Std = proposed standard (0.09 ppm for 1-hour
and 0.070 ppm for 8-hour average), and
OzAttain = rolled-back O_3 value in the "attainment" scenario.

The change in O_3 concentrations is OzCurrent – OzAttain, calculated at the daily level for each site, which is the difference between the observed value and the rolled-back value for each site on each day of the year.

To investigate the sensitivity of Equation (1) to the selection of a *policy-relevant background* level using data from the South Coast Air Basin, varying levels of O_3 background were used to determine the concentrations for the specified percentiles. In 1981,

the daily maximum 1-hour value was 0.427 ppm. Table 2-10 shows the results of applying the rollback methodology using varying levels of *policy-relevant background* O_3 using Equation (1). Note that as the background level increases from 40 ppb, the rate of decline of the daily maximum 1-hour average values hypothesized by the model slows down. Thus, the predicted daily maximum hourly concentrations indicated for each of the percentiles *increases* as the background level increases in the model. The amount of rollback and the rate of change of the concentrations are sensitive to the selection of the *policy-relevant background level* used in Equation (1).

Table 2-10. Applying the rollback methodology using varying levels of policy-relevant background O_3 concentrations for the daily 1-hour maximum. Concentrations in ppb units.

Bkgrnd.	**P40**	**P50**	**P60**	**P70**	**P80**	**P90**	**Max**
	90	113	147	177	217	273	427
40	46.46	49.43	53.82	57.70	62.87	70.10	90
45	50.30	53.01	57.02	60.55	65.26	71.86	90
50	54.24	56.68	60.29	63.47	67.72	73.66	90
55	58.29	60.46	63.66	66.48	70.24	75.51	90
60	62.45	64.33	67.11	69.56	72.83	77.41	90
65	66.73	68.31	70.66	72.73	75.50	79.36	90
70	71.12	72.41	74.31	75.99	78.24	81.37	90

Besides the daily maximum 1-hour average concentration, the selection of a *policy-relevant background* level is critical to the performance of the rollback model using the 8-hour daily maximum concentration. Table 2-11 shows the results of applying the rollback methodology using varying levels of *policy-relevant background* O_3 using Equation (1) with the daily maximum 8-hour concentrations. Note that as the background level increases from 40 ppb, the rate of decline of the daily maximum 8-hour average values hypothesized by model slows down. Thus, the predicted daily 8-hour maximum hourly concentrations for

Table 2-11. Applying the rollback methodology using varying levels of policy-relevant background O_3 concentrations for the daily 8-hour maximum. Concentrations in ppb units.

Bkgrnd.	P40	P50	P60	P70	P80	P90	Max
	68	84	111	135	165	197	294
40	43.31	45.20	48.39	51.22	54.76	58.54	70
45	47.31	48.92	51.63	54.04	57.05	60.26	70
50	51.48	52.79	55.00	56.97	59.43	62.05	70
55	55.82	56.82	58.51	60.02	61.90	63.91	70
60	60.34	61.03	62.18	63.21	64.49	65.85	70
65	65.07	65.41	66.00	66.53	67.18	67.88	70
70	68	70	70	70	70	70	70

each of the percentiles *increases* as the background level increases in the model. Thus, as shown for the daily maximum 1-hour analysis, the amount of rollback and the rate of change of the daily maximum 8-hour concentrations are both sensitive to the selection of the *policy-relevant background level* used in Equation (1).

3. Uncertainties in Using Results from Epidemiology Studies

3.1 Introduction

Concentration-response (CR) functions are equations that relate the change in the number of adverse health effect incidences in a population to a change in pollutant concentration experienced by that population. As pointed out by the Staff, developing concentration-response functions from a vast and not fully consistent literature is a difficult task and ultimately involves subjective evaluations. Staff has provided CR functions for effects of short-term exposure on premature mortality, hospital admissions for respiratory disease, emergency room visits for asthma, school absenteeism, and minor restrictions in activity.

Switzer (2004) has documented that the epidemiologic studies cited by CAAQSOD point to a string of inconsistent results when variations in ambient O_3 are related to variations in mortality, both for short-term and long-term exposures. Examples of the pattern of inconsistent and inconclusive findings include the following:

- Sharply different mortality effect estimates for summer and winter, which should not exist under the model of additive proportional effects that is used in the analyses.
- Instability of O_3 mortality effect estimates resulting from different model specifications of weather effects and time trends.
- Instability of O_3 effect estimates resulting from different selections of monitoring sites within cities.
- Heterogeneity of O_3 effect estimates across cities.
- Ozone effect lags that are inconsistent across cities and across studies.
- Exposure-response relations that are inconsistent across cities and across studies.
- Inconsistencies between short-term and long-term effect studies.

As pointed out by CAAQSOD on page 8-11, there are limitations to epidemiological studies. Firstly, it is not possible to characterize exposure in a precise manner similar to that of a chamber study. Most of the epidemiologic studies rely on regional air pollution monitors, which may not reflect the true exposures at the residences of the study subjects. For O_3 and other gases this may be an issue of significant exposure mismeasurement since some limited evidence suggests a low correlation between personal exposure and ambient concentrations of O_3. In addition, study subjects move around from place to place during the day, so one measurement will not adequately reflect overall exposure. Secondly, epidemiologic studies may be subject to bias from uncontrolled or poorly controlled

confounders, such as seasonality, weather and co-pollutants. Ozone presents a particular challenge because of its seasonal nature and high correlation with temperature.

In addition to the statistical concerns described by CAAQSOD, there are others that have not been adequately addressed by staff. These issues are grouped into the following categories.

1. Confounding of weather and time trends with O_3 effects
2. Heterogeneity of O_3 effects and effect modification
3. Heterogeneity of exposure within study areas
4. The relation between exposure and response
5. Long-term O_3-mortality studies

3.2 Confounding Of Weather And Time Trends With Ozone Effects

Ozone variation is substantially correlated with weather variations. Therefore, special care is needed in separating O_3 effects from the much larger effects of weather. The HEI reanalysis studies point to the sensitivity of O_3 pollutant effect estimates to the precise way in which weather effects are modeled and this is noted as well in CAAQSOD (page 12-68, 12-70). Greater flexibility in modeling weather effects is shown to substantially reduce the apparent pollutant effect estimates. Indeed, as pointed out by Switzer (2004), it is quite possible that the recommended further investigation of weather effects would show that remaining O_3 effects are substituting for the unmodeled weather effects.

While CAAQSOD acknowledges O_3 effect sensitivity to alternative modeling of weather, it does not adequately address a critical modeling assumption -- *additivity* of weather and O_3 effects, an assumption that is built into all the O_3 effect estimates cited by CAAQSOD. The additivity assumption is very strong and it presumes that the incremental effects of O_3 would be the same at any level of temperature and humidity. Thus, for example, the presumption is that incremental O_3 effects are the same at moderate temperature

and humidity as they are at extreme temperature and humidity. If this assumption should fail, then additive modeling of O_3 effects, as relied on by CAAQSOD, can lead to uninterpretable estimates of O_3 effects. This is especially true when effects are not proportionally related to O_3.

CAAQSOD has argued strongly throughout Section 12.4 of the document that only season-specific O_3 effects are meaningful. Underlying the argument is the unstated supposition that O_3 and weather effects are indeed not additive and/or that the community response to O_3 is not proportional to concentration. Additionally, the sharp disagreement between summer and winter finding for O_3 effects argues that the effects of weather may not have been adequately addressed. Shortcomings of existing modeling strategies are made evident by the troubling finding that higher O_3 levels appear to be beneficial on winter days (page 12-71). Without a systematic exploration of weather and O_3 interaction, one cannot conclude that some part of the weather effects is mistakenly attributed to O_3, even for studies that are season specific.

Although there have been a number of studies that attempt to relate O_3 with various health endpoints, CAAQSOD recognizes that the epidemiologic findings are inconsistent with regard to choice of O_3 measure and choice of study area, and have unresolved potential for covariate confounding (page B-21).

3.3 Heterogeneity Of Ozone Effects And Effect Modification

CAAQSOD correctly emphasizes the importance of the 90-city study here in that the same modeling strategy was used for all cities in the study. The reanalysis of this study in HEI (2003) reduced pollutant effect estimates and increased their associated standard error

estimates. The real issue is: when are separate city analyses combinable (as in Figure 12-2) and what is the interpretation of a combined estimate when there are genuine inter-city differences among effect estimates?

The approach used in the NMMAPS multi-city study by Samet. *et al.* (2000) and in their reanalysis HEI (2003) is to allow pollutant effects to be different in different cities and to model this variation as random. Adopting the random effects approach introduces the notion of an overall *population* mean effect. However, this population mean is a model construct which does not reflect the inter-city differences that are part of the model. Formal statistical tests to detect overall heterogeneity among cities will not be informative because of the low power of these tests, as clearly pointed out in the HEI Special Panel review.

It is important to understand the sources of inter-city differences among O_3 effect estimates. *Without a clear understanding, one cannot rule out the possibility that effect estimates are model artifacts.* Therefore, there has been a determined but incomplete effort to relate inter-city effect differences to characteristic differences among cities such as demographics, climate, etc. This is called “effect modification” and is a potentially useful approach. However, Samet *et al.* (2000) could not identify any statistically significant pollutant effect modifiers among those that they examined in their 90-city study. Disparities among cities and different studies could arise, for example, through incorrect treatment of confounding variables or an incorrect characterization of the exposure-response relationship.

Finally, it is important to draw attention to regulatory implications of unresolved differences among O_3 effect estimates for different cities. The regulatory question concerns the implied reduction in health effects that could be expected from a specific regulatory standard. For example, based on results from the multi-city studies, it is reasonable to

suppose that a reduction of ambient O_3 will produce no health benefit in some cities, even based on random-effects models.

3.4 Heterogeneity Of Exposure Within Study Areas

Exposures to ambient O_3 will vary across a community on any given day. CAAQSOD seems not to have taken account of exposure variability in its assessments, although inconsistent findings across different age groups and across seasons (Anderson *et al.*, 1998) suggests that exposure variability can impact estimates of O_3 effects that are based on assumptions that ignore heterogeneity of exposure. There are two sources of population exposure variability for a given ambient concentration. One obvious source is the heterogeneity among individual microenvironmental trajectories, such as variations in time spent outdoors, variations in residential and workplace penetration and air exchange factors. A second source of exposure variation is the spatial heterogeneity of O_3 concentrations, which induces different exposures relative to the monitoring site(s) used to measure ambient O_3.

What is really relevant is whether or not reductions in ambient monitor-site O_3 would produce the reductions in community mortality or other health effects that are implied by models that do not consider how exposure is related to measured ambient O_3. Where multiple O_3 monitoring sites are available, it would be a salutary exercise to compare effects estimated using different monitoring sites. If one monitor records proportionally lower ambient O_3 than a second monitor, the first monitor will show a correspondingly larger unit O_3 effect because both monitors are used to explain the *same* community-wide time series of health effects.

If the concentration-response relationship were exactly linear, and if the population average exposure to ambient O_3 was in constant proportion to the reported ambient O_3, then it could be argued that the estimated effect per unit increase in ambient O_3 is not affected by population variability in exposure. But it is important to distinguish between the unit effects of ambient O_3 and the unit effects of O_3 exposure. The proportionality factor relating population exposures to reported ambient O_3 is likely to be different in different cities. The model-estimated unit effects of ambient O_3 would not then be comparable across cities without an understanding of city-specific relationships between exposures and reported O_3. Combined O_3 effect estimates across cities, such as those reported by NMMAPS, implicitly and implausibly assume that the relation between monitored ambient O_3 and ambient O_3 exposure is the same across cities.

If cross-correlations among monitors are indeed high, and average population exposure is indeed approximately proportional to the monitored O_3 values, and the exposure-response relation is indeed linear, then O_3 effect estimates should be about the same using any standardized combination of monitors to represent exposure. If this is empirically contradicted, then exposure-response may not be linear, monitored O_3 poorly represents population exposure, and the effects of O_3 reductions would be hard to anticipate based on the kinds of models and assumptions that are relied on by CAAQSOD. Studies that try to relate ambient O_3 with personal exposure show inconsistent results as noted by CAAQSOD (B-14).

3.5 The Relation Between Exposure And Response – Non-Linear Relationship

On page 10-13, Staff has indicated that to the extent that there is a population threshold, its use of no threshold modelling approach may not be appropriate. On the other hand, Staff argues that if a threshold model were imposed on the data, it would likely result in a higher estimated beta coefficient or slope for concentrations above the threshold, which would increase the impact per ppb for concentrations above the proposed standard. However, as indicated in this section, there are important considerations that go well beyond Staff's statement above concerning the use of a non-linear model.

As discussed by Staff, assumptions regarding the appropriateness of applying thresholds, and at what level, can have a major effect on health effects estimates. One important issue in estimating O_3 health effects is whether it is valid to apply the CR functions throughout the range of predicted changes in ambient concentrations, even changes occurring at levels approaching the natural background concentration (without any human activity). Staff believes that there is no clear threshold for effects that has been reported, although Staff notes that the issue has not been fully investigated except with reference to ER visits for asthma. Most of the epidemiologic studies used in Staff's estimates have used a log-linear model to represent the relationship between O_3 exposure and the health endpoint. In this case, the relationship between O_3 levels and the natural logarithm of the health effect is estimated by a linear regression.

A linear [proportional] exposure-response relationship is key to many of the inferences and conclusions that CAAQSOD draws from the studies that it has reviewed. For example, the combined analyses in the multi-site studies rely on a proportionality that relates O_3 to mortality that is common to all cities. Studies of exposure error, such as the one by

Zeger *et al*. (2000) which concludes that exposure error may bias effect estimates downward, also rely strongly on the proportional effect hypothesis. Furthermore, the rationale for combining multiple monitors within a city also relies on the assumed proportionality.

Some studies of O_3 health effects have indeed discerned a community-level non-linear exposure-response relationship (Burnett *et al*. 1997, Burnett *et al*. 2001, Steib *et al*. 1996, and Ponce de Leon *et al*. 1996). Additionally, clinical studies with controlled exposure have shown nonlinear relations between exposure and response (e.g., Hazucha *et al*., 1992; Adams, 2003). Exposure measurement error will tend to flatten a non-linear exposure-response curve (Cakmak *et al*., 1999) making it harder to distinguish between linear and non-linear associations. Although there are simulation study reports that specific threshold exposure-response models for a population could be distinguished even in the presence of exposure measurement error, it is not clear to what extent their findings could be generalized.

When non-proportional effects are allowed in the effect estimation model, then estimated pollutant effects have been seen to depart from proportionality, as was seen in many cities in the multi-city analysis by Daniels *et al* (2000), Dominici *et al.* (2002) and Moolgovkar (2003). In these studies, the response is modeled as a low-order parametric spline function of ambient pollution. Application of the spline response model to different cities yielded a variety of response shapes, often with inadequate precision, suggesting that there are statistical difficulties distinguishing between linearity and non-linear spline models. Formal tests for response-function linearity will typically have low statistical power against plausible non-linear alternatives although the implications on non-linearity for regulatory purposes are important. Opting for a linear model because of low statistical power can result in regulatory decisions that will not produce the desired mitigation of health effects.

Unfortunately, CAAQSOD has not addressed the critical issue of exposure-response modeling.

In the multi-city studies, non-linear exposure-response functions for different cities were pooled across cities, as in Schwartz and Zanobetti (2000) and Daniels *et al.* (2000), even though city-to-city differences among pollutant-effect response functions are not obviously within the range of sampling variability. However, a pooled response function, even if it is linear, is not interpretable unless exactly the same exposure-response applies to every city. The putative benefits of ambient pollutant reductions in any particular city cannot be deduced from the pooled response function. However, the conclusions of CAAQSOD rely strongly on questionable commonality and linearity of the pollutant-effect response function, even to the extent that proportionality constants in linear models would need to be the same for every city.

The assumed linearity (proportionality) between O_3 exposure and response has far-reaching regulatory implications. For example, one could double the health effect improvement by doubling the O_3 reduction, so there is no obvious regulatory threshold based on health effects under this assumption. Furthermore, the same reduction of effects could then be achieved by the same reductions in either a high-O_3 or a low-O_3 city.

As pointed out by Switzer (2004), there are also fundamental issues related to the compatibility of community linear exposure-response relationships in the context of individual variations in susceptibility, apart from issues related to individual heterogeneity of exposure. For example, suppose that there are individual-specific response thresholds to 1-hour O_3 exposures. Then the assumed linearity in the community response to the O_3 index implies that the population proportion that responds to pollutant levels that are between 0 and

x is the same as the population proportion that responds to levels between *x* and 2*x*. This is incompatible with plausible assumptions about the population distribution of susceptibility thresholds. Thus, one might question whether the assumed linearity of community exposure-response is even a plausible working hypothesis.

3.6 Acute (Time Series) Mortality Studies

Staff has stated that current epidemiological studies support the preliminary conclusion that warm season O_3 concentrations represent an independent risk factor for mortality (page8-14). The NMMAPS study is cited as the most robust study of O_3 and mortality. Specifically, the reanalysis by Dominici *et al*. (2003) found a positive and significant O_3 effect for the summer period and an equally negative and significant effect for winter (page 12-72). For O_3 to have both a positive and negative effect indicates its effect, if any, is confounded or the model used to estimate its effect is inadequate. In either case, this robust study of 90 individual cities lacks the scientific rigor to support a causal relationship between O_3 and acute mortality, let alone estimating beneficial impacts that should result from controlling ambient O_3 concentrations. In assuming that a cause-effect relationship exists, Staff is ignoring its own advice that (1) additional studies are needed to assure that acute mortality effects are not artifacts arising from confounding by other pollutants, temperature, or weather, misclassification of personal exposure vis-a-vis ambient concentrations, and (2) models have undergone sensitivity analyses (page 8-14).

3.7 Long-Term Mortality Studies

CAAQSOD refers to several long-term ecological cohort studies of O_3 health effects, of which Pope *et al.* (2002) is the latest and most comprehensive. In these long-term studies, both the pollutant concentration and mortality for each city are represented by single average numbers that do not vary over time. The ecological studies cited by CAAQSOD are cohort studies limited to enrolled individuals for whom individual covariate information is available, such as demographic information and smoking habits. The individual covariate information is used to adjust crude mortality rates for the enrolled cohort so as to even out the mortality comparisons between cities. Ozone health effects are inferred by relating time-averaged adjusted mortality to time-averaged monitored O_3 across cities.

Both the Pope *et al.* (2002) cohort study of long-term pollutant effects and the Dominici *et al.* (2002) time-series study of short-term pollutant effects involve a comparable number of U.S. cities. However, geographic variation in the cohort studies takes the place of time variation in the time-series studies. City-specific effect modifiers in time-series studies, as discussed earlier, become confounding variables in the cohort studies. A putative confounding variable in a cohort study is one that shows geographic covariability with PM. Thus, demographic adjustments in the cohort studies are a way of accounting for potential confounding of O_3 effects by demographic variables. Similarly, between-city variations of co-pollutants and climate variables could be related to between-city variations of O_3 and thereby contribute to confounding of O_3 effects.

In some important ways, however, a multi-city cohort study suffers from disadvantages vis-à-vis a single-city time-series study. For example, in a time-series study,

the population at risk is the same each day, while in the cohort study, the population at risk in each city is different and models are needed to bring the separate at-risk populations into alignment. Also, the assigned O_3 concentration for a city needs to be related not to the average city-wide exposure but rather to the average exposure of the cohort assigned to that city. Further, it is reasonable to suppose that this exposure measurement error will be different for different cities.

In any event, the large study by Pope *et al.* (2002) did not discern an O_3 effect on total mortality even when restricted to summer O_3 concentrations and to specific causes of death. The negative findings are noted by CAAQSOD (page 12-51, B-18).

3.8 Staff's Concerns About the Use of Epidemiological Results

There are a variety of unresolved statistical issues in CAAQSOD. Several challenges and unresolved issues present themselves with respect to designing and interpreting time-series studies of O_3-related health effects. As noted by Staff in Section 10.6 (beginning on page 10-11), the principal challenge facing the analyst in the daily time series context is to remove bias due to confounding by short-term temporal factors operating over time scales from days to seasons. The correlation of O_3 with these confounding terms tends to be higher than that for PM or other gaseous pollutants. Thus, model specifications that may be appropriate for PM, the primary focus of much of the available literature, may not necessarily be adequate for O_3. Staff notes that few studies to date have thoroughly investigated these potential effects with reference to O_3, introducing an element of uncertainty into the health benefits analysis.

Of particular importance is the strong seasonal cycle for O_3, high in summer and low in winter, which is opposite to the usual cycle in daily mortality and morbidity, which is high in winter and low in summer. Inadequate control for seasonal patterns in time-series analyses leads to biased effect estimates. In the case of O_3, inadequate seasonal pattern control generally yields statistically significant inverse associations between O_3 and health outcomes. Also, temporal cycles in daily hospital admissions or emergency room visits are often considerably more episodic and variable than is usually the case for daily mortality. As a result, smoothing functions that have been developed and tuned for analyses of daily mortality data may not work as well for removing cyclic patterns from morbidity analyses.

Potential confounding by daily variations in co-pollutants and weather is another analytical issue highlighted by Staff. With respect to co-pollutants, daily variations in O_3 tend not to correlate highly with most other criteria pollutants (e.g., CO, NO_2, SO_2, PM_{10}), but may be more correlated with secondary fine particulate matter (e.g., $PM_{2.5}$) measured during the summer months. Assessing the independent health effects of two pollutants that are somewhat correlated over time is problematic.

Staff notes that the choice of the studies and concentration-response functions used for health impact assessment can affect the benefits estimates. Because of differences, likely related to study location, subject population, study size and duration, and analytical methods, effect estimates differ somewhat between studies. Although the Staff believes that it has addressed this issue by emphasizing meta-analyses and multi-city studies, and also by presenting estimates derived from several studies, Switzer (2004) has noted that serious deficiencies associated with the time-series methodology exist and lead to large uncertainties with the effects estimates.

Another issue relates to the shape of the CR function and whether there is an effect threshold. An important consideration in determining if a safe level of O_3 can be identified is whether the CR relationship is linear across the full concentration range or instead shows evidence of a threshold. Among the O_3 epidemiology literature, only a few studies of hospital admissions and emergency room visits have examined the shape of the CR function. Since only a few studies have investigated whether there is an effect threshold, and the few studies available do not cover all endpoints, Staff notes that the epidemiologic literature does not provide a basis for concluding whether or not there is a population effect threshold. Staff has assumed that there is no threshold for O_3 effects and to the extent that there may not be health effects below the proposed O_3 standard, the analysis may overestimate the impacts of reducing O_3. Staff notes that for the purposes of its analysis, they have estimated benefits down to a background concentration of 0.04 ppm, except for emergency room visits for asthma for which a higher threshold value was used. It is important to note that about 76 - 86% of the benefits presented in Chapter 10 accrue at O_3 concentrations between the proposed standards and background. Staff notes (page 10-13) that to the extent that there is a population threshold, the estimated benefits may not be accurate.

A related issue is that limited data suggest that O_3 effects may be seasonal. While analysis of year round data suggests positive associations between a number of endpoints and O_3 exposure, some data sets that have been analyzed seasonally report positive RR estimates for summer and negative RR estimates for winter. Staff notes that the cause of this has not been adequately investigated, and speculates that this observation may be related to thresholds, differences in personal exposure between seasons, or to co-pollutant exposures.

Staff notes in Chapter 10 that there are likely uncertainties in the statewide exposure assessment, and in whether the existing ambient monitoring network provides representative estimates of exposure for the general population. The available epidemiological studies have used multiple pollutant averaging times, and Staff has proposed conversion ratios for 1-hour to 8-hour and 24-hour O_3 concentrations based on national estimates. Uncertainty is added to the estimated benefits of attainment of the proposed standards to the extent the converted concentration bases differ from monitored concentrations.

3.9 Conclusions

The CAAQSOD synthesis of epidemiologic studies often contains important caveats regarding modeling issues, since effect estimates will be strongly dependent on modeling. However, its own caveats are put aside in drawing conclusions, especially in estimating the benefits described in Chapter 10. The available epidemiologic evidence on O_3 mortality cannot be used to draw robust conclusions regarding the circumstances and magnitudes of ambient O_3 mortality, in particular whether reported O_3 effects are causative. The inconsistencies and model dependency of effect estimates seem to have been brushed aside by CAAQSOD. Without a clear understanding of the reasons for inconsistent effects estimates, one cannot rule out the possibility that O_3 effect estimates are model artifacts. Indeed, CAAQSOD does itself regard the estimates of O_3 mortality effects skeptically (page B-25). Based on a review of the literature, the following conclusions are made:

1. **Sensitivity of O_3 effect estimates to model specification.** This issue was brought to light in the HEI reanalysis in the context of time and weather adjustments, and serves as a cautionary tale. The reported effects of O_3 are often difficult to discern and are inconsistent among cities, regions, seasons, and time lags. Such inconsistencies may be suggestive of modeling inadequacies, particularly in regard to unmodeled confounding and unexplained effect modifiers.

That O_3 effect estimates are delicate is not surprising given that they are superimposed on much stronger effects due to concomitant weather variations, for example. Without a clear understanding of the reasons for inconsistent effects estimates, we cannot rule out the possibility that O_3 effect estimates are model artifacts.

2. **Enforced additivity in the analysis model**. The analysis models relied on by CAAQSOD assume that O_3 effects are necessarily the same at any temperature, even when restricted to summer data. Approaches to mitigate the problem, depending on availability of data, include joint response surface modeling of O_3 and its confounders or stratification of the analyses based on confounder categories.

3. **Enforced linearity of exposure-response**. Because O_3 health effect estimates are inconsistent across studies, cities, seasons, etc., putative benefits of ambient O_3 mitigation are difficult to know. Enforced model linearity of exposure-response, as in the case of the analysis models the CAAQSOD relies on, conceals heterogeneity of response. Pooling of response functions to obtain linearity is not statistically justified and leads to regulatory dilemmas.

4. **Spatial variability of O_3 health effect estimates within cities.** There has been insufficient attention to the issue of spatial variability of effect estimates within cities based on selection or combination of monitors.

5. **Incomplete characterization of the relations between ambient O_3 exposure, individual PM exposure, individual PM susceptibility to health effects, and community level health effect measures**. The models that CAAQSOD uses for the analysis of community health effects of O_3 do not have any link to individual response functions.

6. **Unresolved inconsistencies of O_3 effect estimates.** The following inconsistencies are unresolved: seasonal differences, regional grouping, spatial heterogeneity both between cities and within cities, time lag selection, and treatment of gaseous pollutant confounders.

An alternative to the cause-and-effect explanation provided in CAAQSOD is that many of the results cited for the epidemiology studies may be mostly associated with modelling artifacts. Ultimately the question that must be answered is this: If the time-series data are the most important information available for establishing benefits estimates, are the data good enough to use in the decision-making process? Based on the time-series evidence

presented in the CAAQSOD, one simply cannot draw comfortable conclusions regarding the circumstances and magnitudes of ambient O_3 health effects, or whether reported O_3 health effects are causative. Thus, one might conclude that there is still too much uncertainty remaining in the epidemiological time-series results. In addition, many of the concerns expressed by Staff about the strengths and limitations of the extensive body of epidemiologic evidence of associations between health effects and air pollutants have not been adequately addressed in CAAQSOD. The growing pattern of inconsistent and inconclusive findings using time-series data is troublesome and presents both scientists and policymakers with a very difficult decision. Simply stated, the science is not yet substantial enough, based on epidemiological data, to provide the foundation upon which one can base a health benefits analysis.

4. References

Adams, W. C. (2003) Comparison of chamber and face mask 6.6-hour exposure to 0.08 ppm ozone via square-wave and triangular profiles on pulmonary responses. Inhal Toxicol 15:265-81.

Anderson, H. R.; Ponce de Leon, A.; Bland, J. M.; Bower J. S.; Emberlin, J.; Strachan, D. P. (1998) Air pollution, pollens, and daily admissions for asthma in London 1987-92. Thorax 53:842-8.

Burnett, R. T.; Cakmak, S.; Brook, J. R.; Krewski, D. (1997) The role of particulate size and chemistry in the association between summertime ambient air pollution and hospitalization for cardiorespiratory diseases. Environ Health Perspect 105:614-20.

Burnett, R. T.; Smith-Doiron, M.; Stieb, D.; Raizenne, M. E.; Brook, J. R.; Dales, R. E.; Leech, J. A.; Cakmak, S.; Krewski, D. (2001) Association between ozone and hospitalization for acute respiratory diseases in children less than 2 years of age. Am J Epidemiol 153:444-52.

Cakmak, S.; Burnett, R. T.; Krewski, D. (1999) Methods for detecting and estimating population threshold concentrations for air pollution-related mortality with exposure measurement error. Risk Anal. 19: 487-496.

Daniels, M.; Dominici, F.; Samet, J. M.; Zeger, S. L. (2000) Estimating particulate matter-mortality dose-response curves and threshold levels: an analysis of daily time-series for the 20 largest US cities. Am. J. Epidemiol. 152: 397-406.

Dominici, F.; Daniels, M.; Zeger, S.L.; Samet, J.M. (2002) Air pollution and mortality: estimating regional and national dose-response relationships. J. Am. Statist. Assn. 97: 100-111.

Dominici, F.; Daniels, M.; McDermott, A.; Zeger, S. L.; Samet, J. M. (2003) Shape of the Exposure-Response Relation and Mortality Displacement in the NMMAPS Database. Health Effects Institute Special Report. Health Effects Institute Special Report:91-6.

Fiore, A. M.; Jacob, D. J.; Bey, I.; Yantosca, R. M.; Field, B. D.; Fusco, A. C.; Wilkinson, J. G. (2002). Background ozone over the United States in summer: Origin, trend, and contribution to pollution episodes. J. Geophys. Res. 107(D15): ACH 11-1 – 25.

Fiore, A.; Jacob, D. J.; Liu, H.; Yantosca, R. M.; Fairlie, T. D.; Li, Q. (2003) Variability in surface ozone background over the United States: implications for air quality policy. J. Geophys. Res. (Atmos.): 108: D24, 4787,ACH 19-1 – 19-10.

Hazucha, M. J.; Folinsbee, L. J.; Seal, E. Jr. (1992) Effects of steady-state and variable ozone concentration profiles on pulmonary function. Am Rev Respir Dis 146:1487-93.

Health Effects Institute (2003) Revised analyses of the National Morbidity, Mortality, and Air Pollution Study (NMMAPS). Special report. Boston, MA.

Lefohn, A. S.; Shadwick, D. S.; Ziman, S. D. (1998) The Difficult Challenge of Attaining EPA's New Ozone Standard. Environmental Science & Technology., 32(11):276A-282A.

Moolgovkar, S. H. (2003) Air pollution and daily mortality in two U.S. counties: season-specific analyses and exposure-response relationships. Inhalation Toxicology. 15:877:907.

Ponce de Leon, A.; Anderson, H. R.; Bland, J. M.; Strachan, D. P.; Bower, J. (1996) Effects of air pollution on daily hospital admissions for respiratory disease in London between 1987-88 and 1991-92. J Epidemiol Community Health 50 Suppl 1:s63-70.

Pope, C.A.III; Burnett, R.T.; Thun, M.J.; Calle, et al. (2002) Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. J. Am. Medical Assn. 287: 1132-1141.

Reynolds, S. D.; Blanchard, C. L.; Ziman, S. D. (2003) Understanding the effectiveness of precursor reductions in lowering 8-hr ozone concentrations. J. Air & Waste Manage. Assoc. 53: 195-205.

Reynolds, S. D.; Blanchard, C. L.; Ziman, S. D. (2004) Understanding the effectiveness of precursor reductions in lowering 8-hr ozone concentrations-Part II The eastern United States. J Air & Waste Manage. Assoc. in press.

Samet, J. M.; Dominici, F.; Zeger, S. L.; Schwartz, J.; Dockery, D. W (2000) National morbidity, mortality, and air pollution study. Cambridge, MA: Health Effects Institute; research report no. 94.

Schwartz, J.; Zanobetti, A. (2000) Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily death. Epidemiology 11: 666-672.

Stieb, D. M.; Burnett, R. T.; Beveridge, R. C.; Brook, J. R. (1996) Association between ozone and asthma emergency department visits in Saint John, New Brunswick, Canada. Environ Health Perspect 104:1354-60.

Switzer, P. (2004). A Review of Statistical Methods Used in Epidemiologic Studies of Ambient Ozone and Mortality Chapter 12 and Appendix B. Report prepared for American Petroleum Institute, Washington, DC and Western States Petroleum Association, Sacramento, CA.

U.S. EPA (2003). The Ozone Report - Measuring Progress Through 2003. U.S. Environmental Protection Agency Office of Air Quality Planning and Standards Emissions, Monitoring, and Analysis Division. Research Triangle Park, North Carolina. Report No. EPA 454/ K-04-001.

Winner, D. A.; Cass, G. R. (2000) Effect of emissions control on the long-term frequency distribution of regional ozone concentrations. Environmental Science & Technology. 34: 2612-2617.

Zeger, S. L.; Thomas, D.; Dominici, F.; Samet, J. M.; Schwartz, J.; Dockery, D.; Cohen, A. (2000) Exposure measurement error in time-series studies of air pollution: concepts and consequences. Environ. Health Perspect. 108: 419-426.

ENVIRON

September 24, 2004

MEMORANDUM

To: Kyle Isakower, American Petroleum Institute

From: Stan R. Hayes, ENVIRON

Subject: Initial Comments on Chapter 10 of California's Draft Ozone Staff Report

As a part of their review of the California ambient air quality standard for ozone, on June 21, 2004, the California Air Resources Boards (ARB) and the Office of Environmental Health Hazard Assessment (OEHHA) released for public review and comment a Draft Ozone Staff Report (Cal/EPA 2004a) entitled "Review of the California Ambient Air Quality Standard for Ozone." Subsequently, on August 24, 2004, ARB and OEHHA released Chapter 10 of the draft staff report, entitled "Quantifying the Health Benefits of Reducing Ozone Exposure" (Cal/EPA 2004b).

At the request of the American Petroleum Institute (API), we previously reviewed and prepared comments (Hayes 2004) on portions of the Draft Ozone Staff Report relating to health effects and epidemiology (primarily in Chapters 11 and 12). API has also requested that we review the draft of Chapter 10. This memorandum summarizes our initial comments on that chapter.

The amount of material considered in Chapter 10 and cited by it is large. Given the very limited time and resources available to us, it is not possible at this time to review all of the topics and material cited in the chapter. Therefore, we focus primarily here on the chapter's characterization of ozone concentration-response (CR) relationships. While we also have comments concerning the chapter's assumptions about the level of background ozone and the validity of the proportional rollback procedure used to estimate attainment air quality, we understand that other reviewers are addressing those topics.

As with our earlier comments, we ask that our comments on Chapter 10 be viewed as preliminary. We reserve the right to expand and supplement our comments later, if appropriate and should additional opportunity, time, and resources become available.

SUMMARY

Specific comments are summarized as follows:

1. The validity of concentration-response (CR) function models assumed in past epidemiological studies for assessing ozone's effect on acute mortality and other, non-mortality health endpoints needs to be more rigorously tested, demonstrated, and reconciled with human chamber data.

2. Additional, ozone-specific analyses are needed before drawing conclusions from past epidemiological studies about the effects of ozone on acute mortality and other, non-mortality health endpoints.

3. Quantitative estimation of the incidences of ozone acute mortality in the benefits analysis in Chapter 10 should be deferred until completion of the work recommended in Comments 1 and 2, or at a minimum, treated only through sensitivity analysis.

4. A more comprehensive evaluation and quantification of the nature, magnitude, and implications of uncertainties, individually and in combination, is warranted and needed in the benefits analysis in Chapter 10.

COMMENTS

1. The validity of concentration-response (CR) function models assumed in past epidemiological studies for assessing ozone's effect on acute mortality and other, non-mortality health endpoints needs to be more rigorously tested, demonstrated, and reconciled with human chamber data. For more than twenty years, response to ozone has been directly measured in controlled human chamber studies. Typically, ozone exposures in those studies are well characterized, exercise levels are specified, controls are applied to avoid confounding factors, and response indicators are precisely measured. Collectively, this body of human chamber studies provides a large and robust database with which to characterize human response to ozone over a wide range of ozone levels and lung function, respiratory symptom, and other endpoints.

However, none of these data are used in the benefits assessment in Chapter 10. Instead, Chapter 10 relies entirely on the results of epidemiological studies, bypassing data from human chamber studies. We strongly recommend that Chapter 10 be revised to include those chamber data in its assessment, or at a minimum, to reconcile those data with assumptions made in the chapter based on epidemiological studies.

Given the chapter's current sole reliance on epidemiological studies, it is important that the validity of the concentration-response (CR) function models assumed in those studies for assessing ozone's effect on acute mortality and other, non-mortality health endpoints be more rigorously tested, demonstrated, and reconciled with human chamber data.

While a full analysis of this issue is beyond the scope of this review, we caution that new, ozone-specific analyses to verify the applicability of past epidemiological studies are warranted and necessary. This is particularly the case for the chapter's use of previous PM time-series mortality studies to quantify ozone mortality effects. Absent new ozone-specific analyses showing otherwise, there are reasons for concern about the reliability of using previous epidemiological studies to characterize the effects of ozone exposure on acute mortality and other, non-mortality health endpoints. As noted in our earlier comments, these reasons include the following:

- As described in documentation for USEPA's BenMAP model (Abt 2003), nearly all of the epidemiological studies cited in Chapter 10 assume a log-linear CR function model. This is the case for ozone acute mortality and the other, non-mortality health endpoints considered (a few studies assume a linear or other model for certain non-mortality endpoints). The validity of assuming either a log-linear or a linear model for ozone acute mortality and other, non-mortality health endpoints is not known, and has not yet been tested, demonstrated, or reconciled with human chamber data.

- With respect to the use of PM time-series studies (e.g., NMMAPS as reanalyzed by Dominici et al. 2003) to characterize ozone acute morality effects, concerns about the validity for ozone of the CR function model assumed in those studies include:

  – Ozone presents greater confounding problems than PM, due to its strong seasonal cycle and temperature dependence.

  – The most appropriate metric for ozone exposure is not known and thus may be different than studied in previous PM epidemiological studies. It may need to consider multiple factors such as averaging time (1-hour, 8-hour, or seasonal), different measures of maximum concentration, cumulative dose (concentration x time), inter-episode duration, and frequency of repeated elevated concentrations.

  – Greater uncertainty exists for ozone than for PM about the degree to which fixed-site measurements accurately characterize personal exposure, with the low degree of correlation reported between fixed-site levels and personal ozone exposure a cause for concern about the use of fixed-site data to characterize actual exposure.

  – The effect of the more pronounced diurnal profile for ozone is not known.

As stated earlier, it may be that the first of these reasons – that is, the use of log-linear or linear CR function models in previous epidemiological studies – is of particular importance. Moreover, questions about the validity of log-linear or linear models take on added importance at or below the levels of federal and ozone state standards, since by definition those are the concentrations that would occur upon standard attainment.

As the basis for its assessment of the effects of ozone on premature mortality and other health endpoints, Chapter 10 cites the U.S. Environmental Protection Agency's (USEPA) Section 812 estimates of the health effects associated with implementation of the Clean Air Act Amendments of 1990 (USEPA 1999), the World Health Organization's (WHO) meta-analyses of ozone (Anderson et al. 2004), and the Levy et al. (2001) analysis of the public health benefits of reducing ozone.

We do not have time or resources to review in detail these references and the large number of individual studies that they cite. However, our initial survey of the three references suggests that the cited studies nearly all assume a log-linear CR function model, both for acute mortality and for nearly every other endpoint (see also Abt 2003). We recommend that the evidence for, and reasoning in support of, this important assumption for ozone be added to Chapter 10 and elsewhere in the Draft Ozone Staff Report.

Because of the strength and quality of the large database available from human chamber studies, we recommend that Chapter 10 attempt to reconcile its log-linearity assumption with the results of those studies. While there is no assurance that the presence of absence of log-linearity in those response data would be probative with respect to ozone acute mortality and other health endpoints, if log-linearity were to hold for ozone mortality and other acute health endpoints as assumed in Chapter 10 and elsewhere in the Draft Ozone Staff Report, one might expect (or at least hope) to see signs of that log-linearity (or linearity) in the data for other measures of acute ozone response reported from the human chamber studies, including lung function impairment, lower respiratory symptoms, and others.

We recommend that this issue be examined as closely and resolved as definitively as feasible in a revised version of Chapter 10. While we cannot examine all the data sets and acute health endpoints cited in Chapter 10, in an effort to be helpful, we offer here several examples for purposes of illustration (also included in our previous comments). A log-linearity hypothesis for acute lung function impairment and lower respiratory symptoms can be tested using, for example, human clinical data from controlled chamber studies conducted by Avol et al. (1984), Kulle et al. (1985), and McDonnell et al. (1983). These studies, which are also cited in the Draft Ozone Staff Report, measured the response of heavily exercising, healthy young adults exposed to a range of ozone concentrations for one hour (Avol) or two hours (Kulle and McDonnell). Response to ozone exposure was measured in terms of lung function impairment (e.g., FEV1[1] decrement) and lower respiratory symptom incidence (e.g., cough and chest discomfort).

During an earlier review of the national ambient air quality standards (NAAQS) for ozone, Hayes et al. (1987; 1989) used data from these three studies to develop exposure-response relationships that were cited by USEPA in their Ozone Staff Paper (USEPA 1988) and used in USEPA's Acute Ozone Health Risk Assessment (Hayes et al. 1989; Winkler et al. 1990; Whitfield et al. 1994).

A full set of exposure-response relationships from Hayes et al. (1989) was provided in our previous comments. From that full set, exposure-response relationships for two health endpoints have been selected. Ozone-induced responses, expressed as the fraction of the healthy adult population fraction affected, are shown in Figure 1 for FEV1 decrements $10% and in Figure 2 for chest discomfort (lower respiratory symptoms for Avol) of any severity (mild, moderate, or severe).

Data for FEV1 decrement responses $10% from the three studies in Figure 1 are re-plotted in Figure 3 on the same graph (upper) and as the natural logarithms of the responses (lower graph). Neither the responses themselves nor their natural logarithms exhibit the log-linear or linear shape assumed in epidemiological studies. A similar result is shown in Figure 4 for "any" chest discomfort. Though not shown, similar results are found for the other exposure-response relationships.

These exposure-response relationships do not exhibit the hypothesized log-linearity (or linearity, either). This also appears to be the case for response data from McDonnell and Smith (1994) (see our earlier comments in which FEV1 response data from 1-, 2-, and 6.6-hour exposure are plotted).

The effect of longer exposure times on this question of CR function log-linearity or linearity is not known, particularly at the lower ozone concentrations characteristic of standard attainment. Although it is uncertain how representative the extended, heavy-exercise protocol used is of typical behavior in the general population, data from Folinsbee et al. (1991), as presented in the Draft Ozone Staff Report (Figure 11-2, p. 11-55), are used to derive the relationships in Figure 5, which are for three levels of FEV1 decrement ($10%, $15%, and $20%) under 6.6-hour exposures at exercise levels intended to represent heavy or strenuous work or play over the period.

The narrow range of ozone exposure concentrations considered by Folinsbee et al. (0.08, 0.10, and 0.12 ppm) makes it difficult to assess the degree of log-linearity or linearity of the exposure-

[1] FEV1 is the volume of air that can be expelled in the first second of a maximal expiration.

response relationship, particularly at lower concentrations (where there are no data). The shape of the exposure-response relationship at those lower concentrations is especially important because it is those levels that by definition would occur under conditions of standard attainment. Although not definitive, and less obviously nonlinear than in Figures 1 through 4, the data in Figure 5 are suggestive of possible response nonlinearities (the strength of which might be related to the definition of what is regarded to be an adverse effect, that is, FEV1 decrements $10%, $15%, or $20%). Further investigation is warranted, particularly at concentrations below 0.08 ppm, which is the lower end of the concentration range considered by Folinsbee et al.

While the epidemiological studies cited by Chapter 10 focus on ozone acute effects, some have questioned whether chronic effects occurring early in life might affect mortality and morbidity later in life, perhaps manifesting themselves in what might be detected as an acute effect then. Relevant to that question, Gauderman et al. (2004) published a study in the New England Journal of Medicine on September 9, 2004. That study examined the effects over time of air pollution on growth of lung function in children during the period of rapid lung development from ages 10 through 18. The authors used linear regression to investigate the relationship between air pollution levels and FEV1 and other spirometric measures. They examined the association between the growth of FEV1 during those ages and air pollution levels in twelve Southern California communities. Exposures to ozone, nitrogen dioxide, acid vapor, particulate matter (PM10 and PM2.5), and elemental carbon were considered.

Gauderman et al. concluded in their paper that "the results of this study indicate that current levels of air pollution have chronic, adverse effects on lung development in children from the age of 10 to 18 years, leading to clinically significant deficits in attained FEV1 as children reach adulthood." However, while urging that results for ozone be interpreted cautiously, the authors also stated that "our results provide little evidence that ambient ozone at current levels is associated with chronic deficits in the growth of lung function in children."

The basis for that statement is shown in Figure 6, which is reproduced from their paper. The figure plots the percent of children with FEV1 less than 80% of the predicted value against the time-averaged air quality in the community in which they lived (10 a.m. to 6 p.m. averaged for ozone). The absence of a significant slope to the linear regression line supports the finding of Gauderman et al. that FEV1 deficit was not found to be statistically associated with ozone levels in those communities.

In summary, the evidence discussed here (mainly for lung function and lower respiratory symptoms from human chamber studies) is not probative with respect to questions concerning ozone acute mortality and other health endpoints. Nonetheless, it does suggest that further investigation is warranted. New, ozone-specific research is needed to better and more systematically assess the plausibility, mechanism, likelihood of occurrence, and potential magnitude of ozone mortality and other, non-mortality effects at concentration levels currently found in ambient air, especially at the lower ozone levels characteristic of standard attainment.

2. Additional, ozone-specific analyses are needed before drawing conclusions from past epidemiological studies about the effects of ozone on acute mortality and other, non-mortality health endpoints. As recommended in our previous comments (Hayes 2004), conclusions in the Draft Ozone Staff Report regarding the effect of ozone exposure on acute mortality and other, non-mortality health endpoints, and particularly the quantification of those effects in ozone benefit analyses in Chapter 10, should be deferred until further assessment of the validity of using

epidemiological studies for that purpose. If warranted based on the findings of that assessment, new ozone-specific research should be conducted to more fully investigate the relationship between ozone exposure and acute mortality and with other, non-mortality health endpoints. Again, as stated in our previous comments, research should include at a minimum the following:

- Use of a nonlinear or non-log-linear ozone concentration-response model, one that at least allows for the possibility of a broader range of responses.

- Inclusion of a range of alternative ozone exposure metrics, considering, in addition to measures of the peak concentration level (e.g., 1-, 8-, 24-hr, seasonal averages), the following other potentially important factors: (a) frequency of peaks, (b) accumulated dosage (C x T), and (c) inter-episode duration (or respite).

- Use of improved ozone-specific techniques for filtering out such confounding influences as ozone's strong seasonal and diurnal cycles, and temperature dependence.

- Use of improved methods to ensure accurate quantification of important uncertainties and their standard setting implications.

3. Quantitative estimation of the incidences of ozone acute mortality in the benefits analysis in Chapter 10 should be deferred until completion of the work recommended in Comments 1 and 2, or at a minimum, treated only through sensitivity analysis. For the reasons outlined in Comments 1 and 2, additional ozone-specific work is needed before including ozone acute mortality (and perhaps other, non-mortality health effects as well) in the benefits analyses in Chapter 10. Relevant to the need for new, ozone-specific analyses (at least for mortality), we note the Draft Ozone Staff Report's assessment that:

> "While there is a real potential for the occurrence of these outcomes [mortality], based on the inflammatory response generated from ozone exposure, additional studies need to be conducted to ensure that: (1) ozone is not confounded by other pollutants including particulate matter (PM10 and PM2.5); (2) ozone is not confounded by temperature and season using parametric (versus non-parametric) generalized linear models; and (3) personal exposure to ozone is sufficiently related to ambient concentrations of ozone. Finally, the ozone-specific models need to undergo the thorough sensitivity analysis of their results similar to that undertaken for studies on particulate matter." (Section 8.3.3, p. 8-14)

Until work such as described in Comments 1 and 2 and in the above-referenced text has been performed, the benefits analysis in Chapter 10 should at least defer quantitative estimation of ozone acute mortality. At a minimum, such estimates, if ARB and OEHHA feel that they must be made, should be accompanied by extensive caveats and broader uncertainty evaluation and perhaps treated through sensitivity analyses, as done by USEPA in its Section 812 benefits analyses (USEPA 1999).

4. A more comprehensive evaluation and quantification of the nature, magnitude and implications of uncertainties, individually and in combination, is warranted and needed in the benefits analysis in Chapter 10. Significant uncertainties exist in the available health and epidemiological data, several reasons for which are noted above. A more quantitative and systematic approach to the characterization of uncertainty would assist in judging the standard setting implications of that uncertainty and identifying areas where additional research is most important. Where appropriate (e.g., ozone acute mortality), the uncertainty analysis should

Where appropriate (e.g., ozone acute mortality), the uncertainty analysis should extend beyond statistical confidence intervals about the mean estimates of health effect incidences reported in Chapter 10 to include a broader assessment of the degree of consensus within the scientific community about ozone's role with respect to the health effect at issue, key remaining questions regarding that role, and the further research needed to answer those questions.

With respect to the question of ozone mortality effects, we note that in their Section 812 benefits analyses, USEPA stated: "While the growing body of epidemiological studies suggests that there may be a positive relationship between ozone and premature mortality, there is still substantial uncertainty about this relationship." As noted earlier, USEPA treated ozone mortality effects in that analysis through sensitivity analysis (USEPA 1999).

REFERENCES

Abt. 2003. "BenMAP: Environmental Benefits Mapping and Analysis Program, User's Manual." Abt Associates Inc., Bethesda, Maryland. Prepared for U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina. November.

Anderson, H.R., R.W. Atkinson, J.L. Peacock, L. Marston, and K. Konstantinou. 2004. Meta-analysis of time-series studies and panel studies of particulate matter (PM) and ozone. Report of a WHO task group. World Health Organization.

Avol, E.L., W.S. Linn, T.G. Venet, D.A. Shamoo, and J.D. Hackney. 1984. Comparative respiratory effects of ozone and ambient oxidant pollution exposure during heavy exercise. *J. Air Pollut. Control Assoc.*, 34:804-809.

Cal/EPA. 2004a. "Review of the California Ambient Air Quality Standard For Ozone, Volumes I and 2." Public Review Draft. California Environmental Protection Agency, Air Resources Board and Office of Environmental Health and Hazard Assessment, Sacramento, California. June 21.

Cal/EPA. 2004b. "Chapter 10, Quantifying the Health Benefits of Reducing Ozone Exposure." Public Review Draft. California Environmental Protection Agency, Air Resources Board and Office of Environmental Health and Hazard Assessment, Sacramento, California. August 24.

Dominici. F., A. McDermott, M. Daniels, S.L. Zeger, and J.M. Samet. 2003. "Mortality Among Residents of 90 Cities. Revised Analyses of Time-series Studies of Air Pollution and Health." Health Effects Institute, Boston, Massachusetts.

Folinsbee, L. J., D.H. Horstman, H.R. Kehrl, W.F. McDonnell, T.R. Gerrity, E. Seal, R. Larson, M.J. Hazucha, S. Abdul-Salaam, B. Faucette, and P.J. Ives. 1991. Effects of single and repeated prolonged low-level ozone exposure in man. Presented at: annual meeting of the Society for Occupational and Environmental Health; March; Washington, DC.

Gauderman, W.J., E. Avol, F. Gilliland, H. Vora, D. Thomas, K. Berhane, R. McConnell, N. Kuenzli, F. Lurmann, E. Rappaport, H. Margolis, D. Bates, and J. Peters. 2004. The ef-

fect of air pollution on lung development from 10 to 18 years of age. *N. Engl. J. Med.*, 351:11, pp. 1057-1067. September.

Hayes, S.R., M. Moezzi, T.S. Wallsten, and R.L. Winkler. 1987. "An Analysis of Symptom and Lung Function Data from Several Controlled Ozone Exposure Studies." SYSAPP-86/120. Systems Applications, Inc., San Rafael, California, for the U.S. Environmental Protection Agency, Research Triangle Park, North Carolina.

Hayes, S.R., A.S. Rosenbaum, T.S. Wallsten, R.G. Whitfield, R.G. Whitfield, and R.L. Winkler. 1989. Acute ozone exposure-response relationships for use in health risk assessment. Paper No. MPM-L4. Annual Meeting, Society for Risk Analysis, San Francisco, California. October 29 – November 1.

Hayes, S.R., A.S. Rosenbaum, T.S. Wallsten, R.G. Whitfield, R.L. Winkler, and H. Richmond. 1989. A health risk assessment for use in setting the U.S. primary ozone standard. In: Atmospheric Ozone Research and Its Policy Implications, T. Schneider and S.D. Lee (Editors), Elsevier Science Publishers, Amsterdam, The Netherlands.

Hayes, S.R. 2004. "Initial Comments on California's Draft Ozone Staff Report." Memorandum, dated August 31.

Kulle, T.J., L.R. Sauder, J.R. Hebel, and M.D. Chatham. 1985. Ozone response relationships in healthy non-smokers. *Am. Rev. Respir. Dis.*, 132:36-41.

Levy J.I., T.J. Carrothers, J.T. Tuomisto, J.K. Hammitt, and J.S. Evans. 2001. Assessing the public health benefits of reduced ozone concentrations. *Environ. Health Perspect.* 109:1215-26.

McDonnell, W.F., D.H. Horstman, M.J. Hazucha, E. Seal, Jr., E.D. Haak, S. Salaam, and D.E. House. 1983. Pulmonary effects of ozone exposure during exercise: dose-response characteristics. *J. Appl. Physiol.: Respir. Environ. Exercise Physiol.*, 54:1345-1352.

McDonnell, W.F., and M.V. Smith. 1994. Prediction of acute ozone response as a function of exposure rate and total inhaled dose. *J. Appl. Physiol*. 76: 2776-2784.

Samet, J.M., S.L. Zeger, F. Dominici, F. Curriero, I. Coursac, D.W. Dockerey, J. Schwartz, and A. Zanobetti. 2000. "The National Morbidity, Mortality, and Air Pollution Study, Part II: Morbidity and Mortality from Air Pollution in the United States." HEI Research Report Number 94, Part II. Health Effects Institute, Cambridge, Massachusetts. June.

USEPA. 1988. "Review of the National Ambient Air Quality Standards for Ozone: Preliminary Assessment of Scientific and Technical Information, Staff Paper." U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina.

USEPA. 1996. "Review of the National Ambient Air Quality Standards for Ozone, Assessment of Scientific and Technical Information, OAQPS Staff Paper." EPA-452\R-96-007. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina. June.

USEPA. 1999. The benefits and costs of the clean air act 1990 to 2010: EPA report to Congress. EPA-410-R-99-001. U.S. Environmental Protection Agency, Office of Air and Radiation and Office of Policy. Washington, D.C. November.

Whitfield, R.G., H.M. Richmond, S.R. Hayes, A.S. Rosenbaum, T.S. Wallsten, R.L. Winkler, M.L.G. Absil, and P. Narducci. 1994. *Chapter 5, Health Risk Assessment of Ozone*. In: *Ozone: Effects on Public Health and Ecological Impacts, Why Society Must Control Tropospheric Ozone*, ed., D.J. McKee. Chelsea, Michigan: Lewis Publishers.

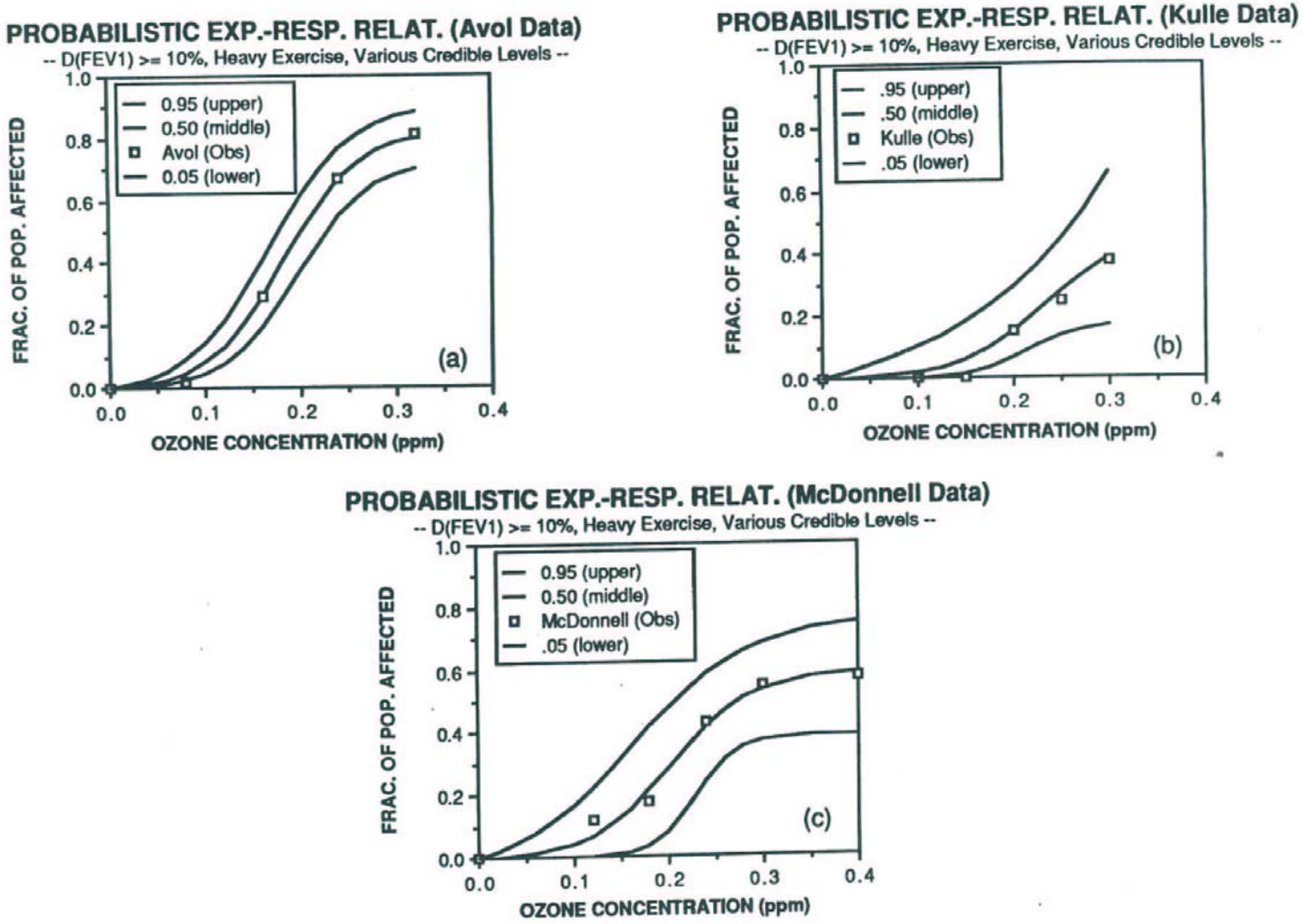


Fig 1. Ozone probabilistic exposure-response relationships for FEV1 decrement ≥10 percent (heavy exercise).

Fig. 1. Figure 1 from Hayes et al. (1989) – FEV1 decrement $ 10% for 1-2 hour exposures to filtered air and ozone at heavy exercise (Avol et al., 1984; Kulle et al. 1985; McDonnell et al. 1983).

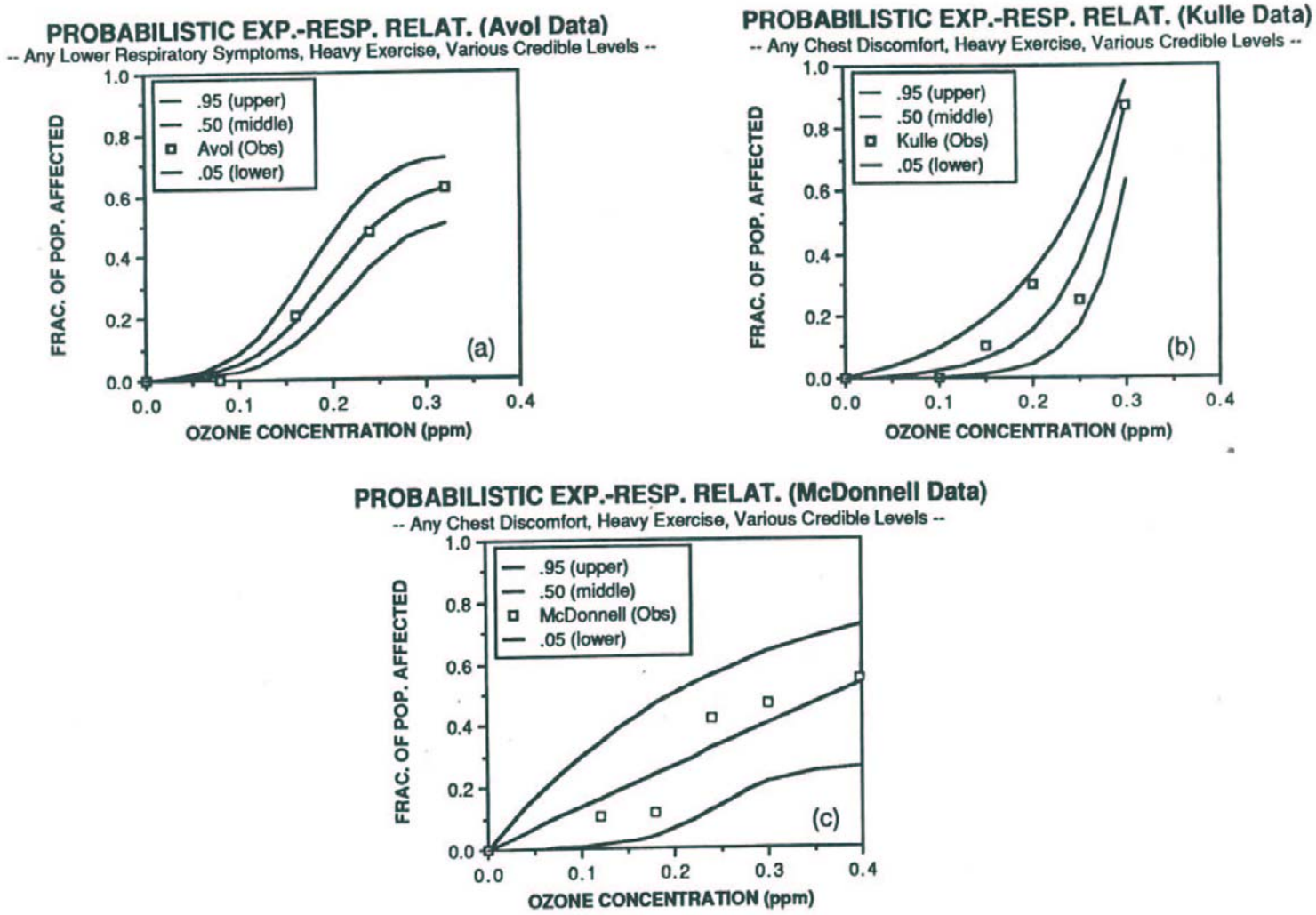
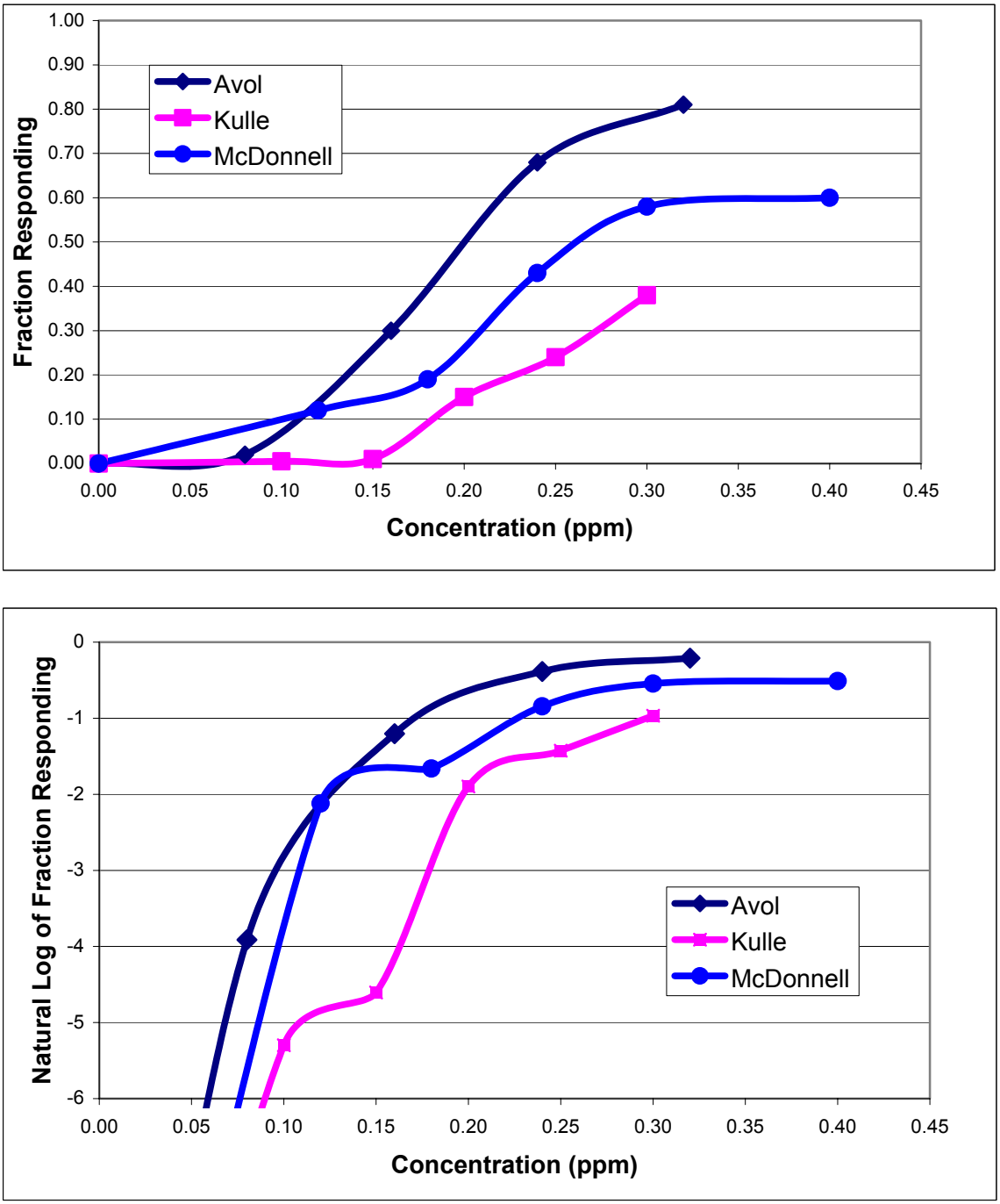


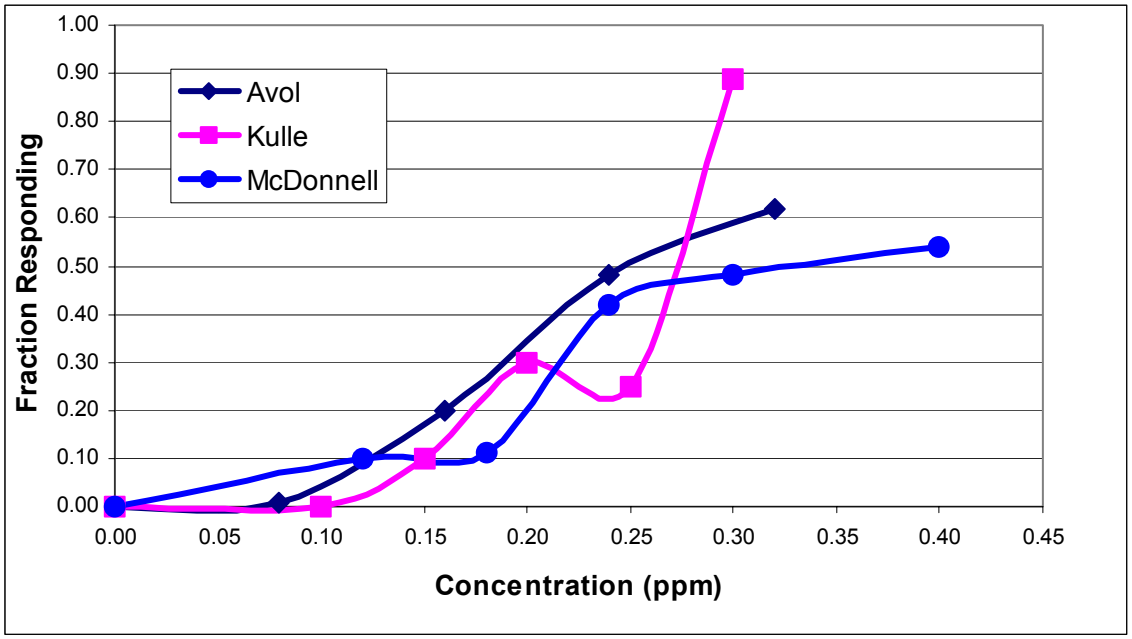
Fig 5. Ozone probabilistic exposure-response relationships for any chest discomfort (mild, moderate, or severe; heavy exercise).

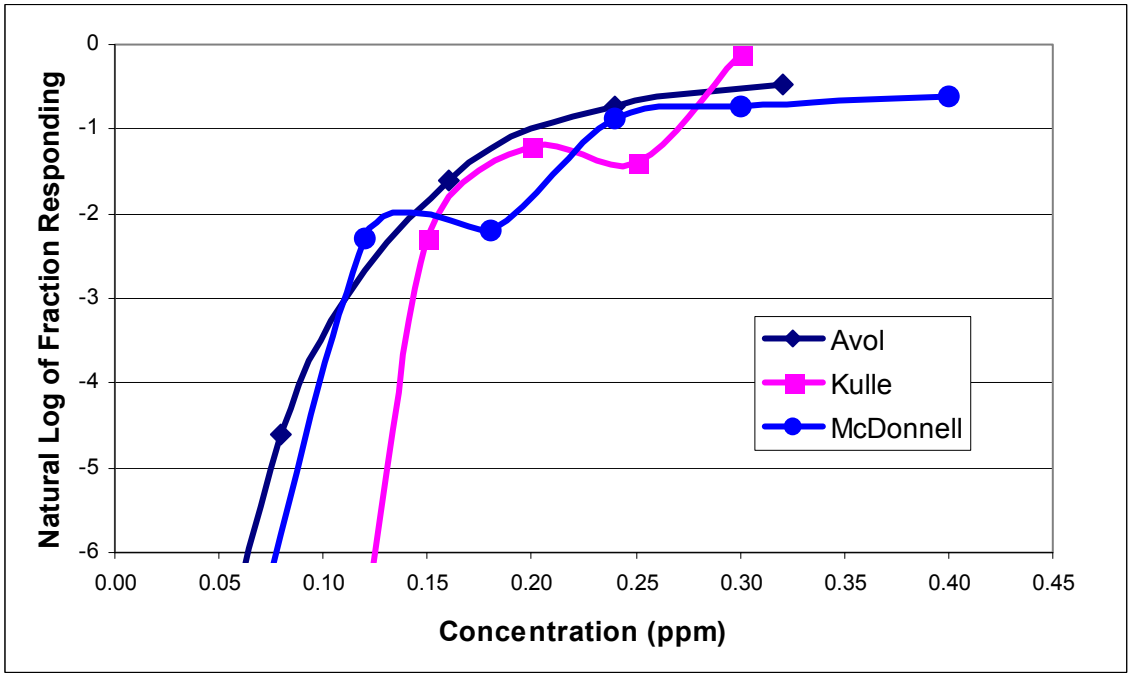
Fig. 2. Figure 5 from Hayes et al. (1989) – lower respiratory symptoms (any chest discomfort) for 1-2 hour exposures to filtered air and ozone at heavy exercise (Avol et al., 1984; Kulle et al. 1985; McDonnell et al. 1983).



Derived from Hayes et al. (1989)

Fig. 3. FEV1 decrement $10% for 1-2 hour exposures to filtered air and ozone at heavy exercise (observations reported by Avol et al., 1984; Kulle et al. 1983; McDonnell et al. 1985) – fraction responding (upper figure) and natural logarithm of fraction responding (lower figure).





Derived from Hayes et al. (1989)

Fig. 4. Lower respiratory symptoms (any chest discomfort) for 1-2 hour exposures to filtered air and ozone at heavy exercise (Avol et al., 1984; Kulle et al. 1985; McDonnell et al. 1983) – fraction responding (upper figure) and natural logarithm of fraction responding (lower figure).

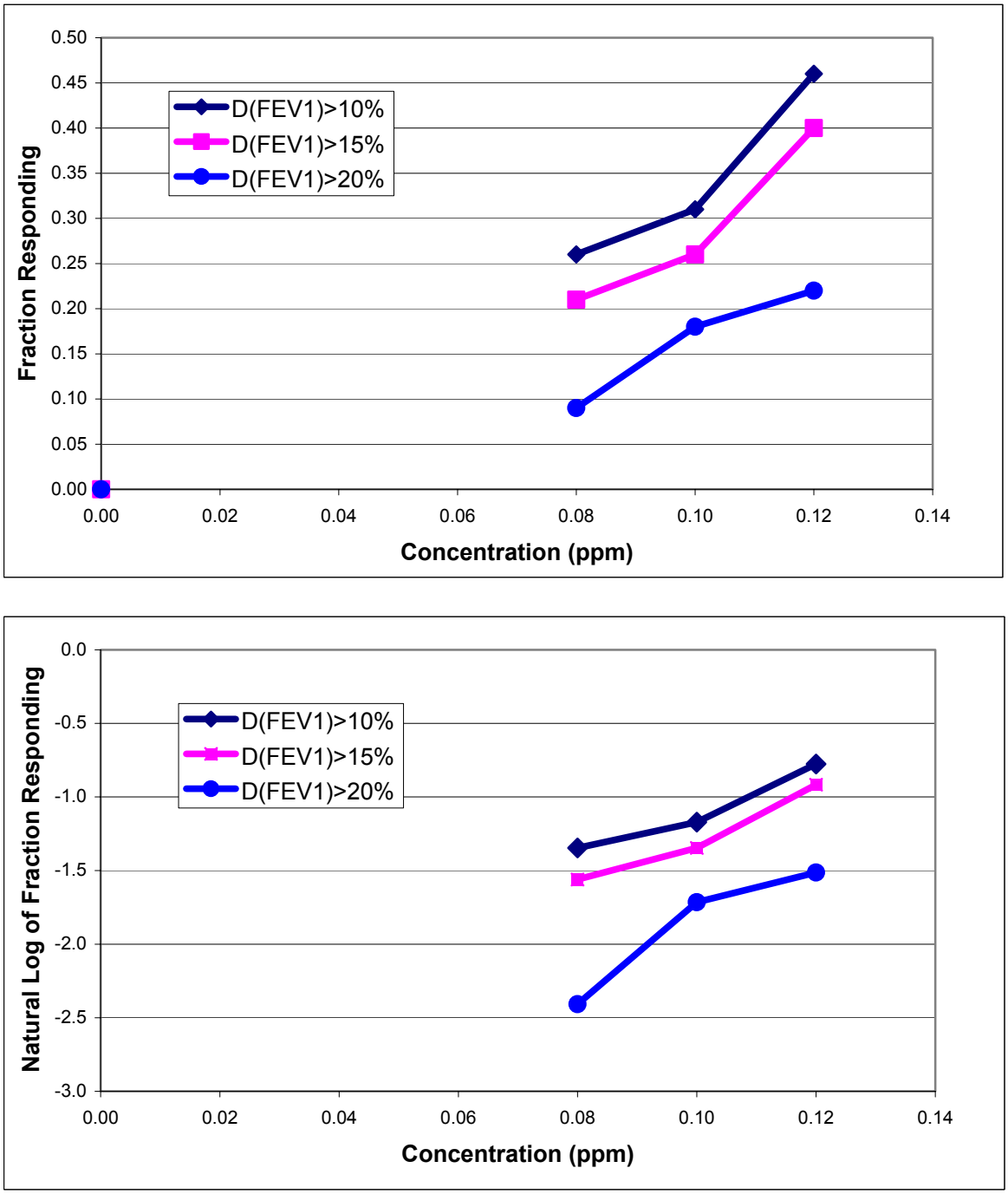
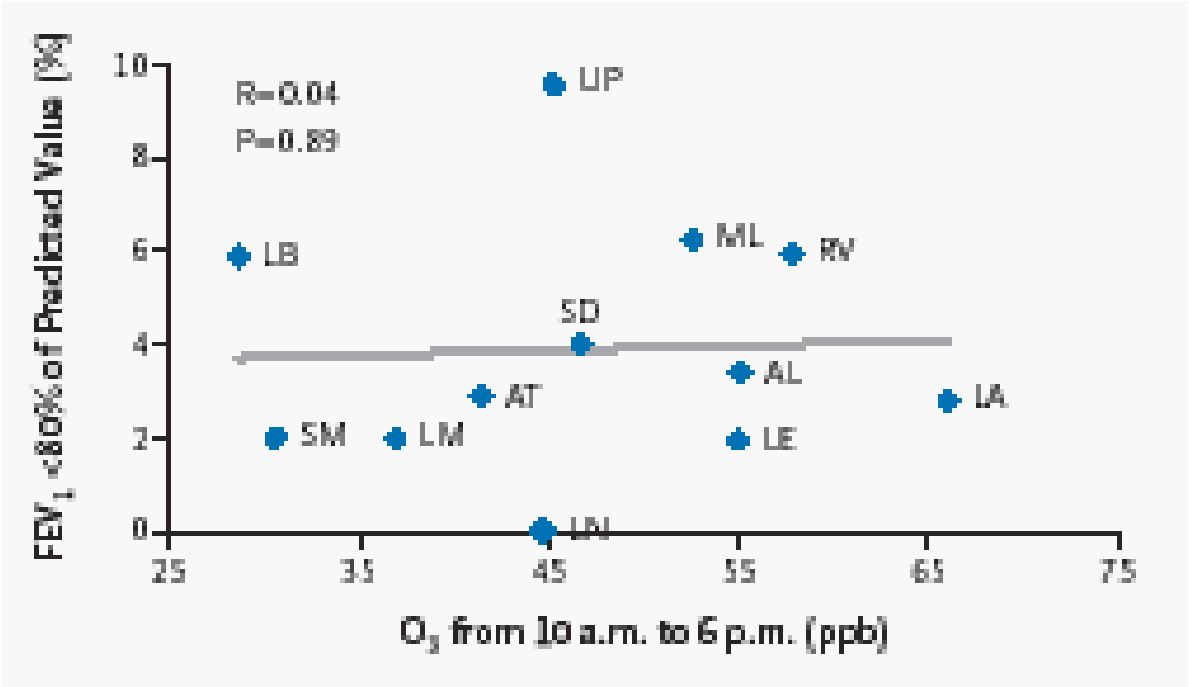


Fig. 5. FEV1 decrements for 6.6 hour exposures to filtered air and ozone at heavy exercise (derived from Follinsbee et al. 1991, as reproduced in Draft Ozone Staff Report, Fig. 11-12, p. 11-55) – fraction responding (upper figure) and natural logarithm of fraction responding (lower figure).



From Gauderman et al. (2004)

Fig. 6. Community-specific proportion of 18-year-olds with a FEV1 below 80 percent of predicted value plotted against the average levels of 10 a.m. – 6 p.m. averaged ozone from 1994 through 2000 in twelve Southern California cities (AL Alpine, AT Atascadero, LE Lake Elsinore, LA Lake Arrowhead, LN Lancaster, LM Lompoc, LB Long Beach, ML Mira Loma, RV Riverside, SD San Dimas, SM Santa Maria, and UP Upland).